EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

21 July 2011
EMA/434565/2010
Committee for Medicinal Products for Human Use (CHMP)

Overview of comments received on 'Guideline on validation of bioanalytical methods' (EMEA/CHMP/EWP/192217/2009)

Interested parties (organisations or individuals) that commented on the draft document as released for consultation.

Stakeholder no.	Name of organisation or individual
1	3S Pharmacological Consultation & Research GmbH
2	American Association of Pharmaceutical Scientists, Bioanalytical Focus Group.
3	American Association of Pharmaceutical Scientists Ligand Binding Assay Bioanalytical Focus Group (AAPS LBABFG)
4	ABS Laboratories Ltd, UK
5	ACC GmbH Analytical Clinical Concepts, Leidersbach, Germany
6	ACRO-CZ, Prague 4, Czech Republic
7	AESGP
8	Algorithme Pharma (Bioanalytical Services Department)
9	Amgen Global Regulatory Affairs and Safety
10	Analytisches Zentrum Biopharm GmbH Berlin, Germany
11	Anapharm Europe
12	Aurigene Discovery Technologies Limited, Hyderabad, India
13	Baxter BioScience
14	Bayer Schering Pharma AG, Drug Metabolism and Pharmacokinetics - Bioanalytics
15	Bioanalytik Forum Bergheim
16	Bristol-Myers Squibb, Co
17	German Pharmaceutical Industry Association BPI e.V, BioManufacturing Working Party
18	CEPHA s.r.o., Pilsen, Czech Republic
19	Developmental Drug Metabolism & Pharmacokinetics, Gedeon Richter Plc.
20	DLR Consulting
21	EFPIA
22	European Generics Association

7 Westferry Circus • Canary Wharf • London E14 4HB • United Kingdom
Telephone +44 (0)20 7418 8400 **Facsimile** +44 (0)20 7418 8416
E-mail info@ema.europa.eu **Website** www.ema.europa.eu

Stakeholder no.	Name of organisation or individual
23	European Quality Assurance Confederation
24	European Bioanalysis Forum (EBF)
25	GE Healthcare Ltd.
26	Gilead Sciences International Limited
27	GlaxoSmithKline Biologicals
28	Guerbet, France
29	GVK Biosciences Pvt Ltd, India
30	Institute für pharmazeutische Chemie, University of Braunschweig, Germany
31	HEXAL AG, Clinical Research Department, Germany
32	Huntingdon Life Sciences
33	IFAPP (International Federation of Associations of Pharmaceutical Physicians)
34	LFB Biotechnologies
35	Macleods Pharmaceutical Limited, Department of Bioequivalence, Mumbai, India
36	Pfizer
37	Pharmaceutical Research Institute (Instytut Farmaceutyczny), Warszawa, Poland
38	Pharmaceutical Research Unit, Amman, Jordania
39	Pharmakl spol. s r.o., Prague, Czech Republic
40	Pharmascience Inc. Canada
41	PRA International, Early Development Services, Bioanalytical Laboratory, Assen, The Netherlands
42	Quest Life Sciences Pvt Ltd, Chennai, India
43	QUINTA-ANALYTICA, s.r.o., Prague, Czech Republic
44	Food and Consumer Product Safety Authority, The Netherlands
45	F. Hoffmann-La Roche Ltd
46	Science Pharma (Poland)
47	Society of Quality Assurance (SQA).
48	Spi-BIO/CEA and Biotec-Centre, France
49	Trident Bioanalytics Ltd.
50	Veeda Clinical Research Pvt. Ltd., Ahmedabad, India
51	Bioanalytical Chemistry, Elusys Therapeutics, Inc., Pine Brook, NJ
52	Plasma Protein Therapeutics Association (PPTA)

1. General comments – overview

Stakeholder no.	General comment (if any)	Outcome (if applicable)
1	This guideline draft is a substantial improvement over the generally accepted FDA 2001 bioanalytical method validations guideline; however a few point need improvement and/or clarifications.	Comment noted.
1	No procedure on extraction recovery evaluation is described in the draft. **Proposed change:** This should be added.	Extraction recovery is an issue investigated during the analytical method development and as such considered not necessary to be included in this guideline.
2	AAPS Bioanalytical Focus Group (BFG) is pleased to have the opportunity to comment on the draft document on bioanalytical methods validation released by EMA. The AAPS BFG membership is made up of AAPS members with an interest in bioanalysis and from a broad cross section of the pharmaceutical industry, CROs and consultants. The BFG is primarily focused on chromatographic methodology but, as will be noted from the feedback, many members are also involved with ligand binding assays. The following are the combined comments from the membership of the BFG. Please note that while comments are separated from suggested changes we ask that all feedback is considered as proposal for changes to the guidance language. While the release of the draft guidance document has heightened the interest in bioanalytical community to update the available guidance on the evolving science of bioanalysis, BFG strongly recommends that EMA, along with other major regulatory agencies in the world, consider joining efforts to create a globally harmonized guidance on bioanalysis, and avoid creating individual guidance (ref: 1-3). Ref 1: Request for global harmonization of the guidance for bioanalytical method validation and sample analysis (An open letter), P. Timmerman, S. Lowes, D. Fast, and F. Garofolo, Bioanalysis (2010) 2(4), 683 Ref 2: International harmonization of bioanalytical guidance (An editorial), S. Bansal, M. Arnold, F. Garofolo, Bioanalysis (2010) 2(4), 685–687 Ref 3: Towards harmonized regulations for bioanalysis: moving	The EMA considers harmonisation an important topic and has therefore embarked into numerous discussions in relation to the preparation of this guideline. The outcome of these interactions on harmonisation issues have been taken into account as much as possible.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	forward! (An editorial), Peter van Amsterdam, Berthold Lausecker, Silke Luedtke, Philip Timmerman, Margarete Brudny-Kloeppel, Bioanalysis, April 2010, Vol. 2, No. 4, Pages 689-691	
3	(AAPS) Ligand Binding Assay Bioanalytical Focus Group (LBABFG) thanks the EMA for providing a guidance document on the Validation of Bioanalytical Methods, and appreciates the opportunity to provide comments on this draft guideline. The consolidated comments of eleven member companies (comprised of Pharmaceutical and Biotechnology companies, Contract Research Organizations and Consultants) are summarized below. Comments from the responding companies were each the result of their own internal collaborative efforts, and therefore reflect the opinions of many interested parties. Overall, the responding companies of the LBABFG strongly support a harmonization effort among all international regulatory agencies. The general consensus of the LBABFG is that this draft guidance does not completely meet the needs of the LBA community. The LBABFG recommends a separation of the two disciplines within a single guidance – one section focused on LC-MS validation and sample analysis and one section focused strictly on LBA validation and sample analysis. A guidance for bioanalytical method validation should take into consideration the similarities and differences between both the method validation requirements for small molecules and large molecules, and the technologies of LC-MS and LBA. A dedicated section for LBA will provide greater clarity to the LBA community for basic validation concepts by providing an environment for a more thorough discussion of the specific elements critical to LBA method validation. In addition to the multiple platforms and assay configurations at our disposal, such concepts include lack of a homogeneous reference standard, calibration curve fitting and the standard curve's inherent limitations, prozone (hook) effect, selectivity and specificity, specific acceptance criteria and critical reagents. Other elements that impact large molecule method development include the degradation of the molecule by enzyme digestion, rather than classical metabolism, and the use of intact matrix. Distinct guidance will also provide greater clarity to the LC-MS	The EMA considers harmonisation an important topic and has therefore embarked into numerous discussions in relation to the preparation of this guideline. The outcome of these interactions on harmonisation issues have been taken into account as much as possible. - It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	community for basic validation concepts that are specific and critical to the LC-MS technology. Such concepts include use of internal standards, metabolite testing, selectivity, use of multiple analytes, carry-over, and use of single sample analysis. More comprehensive comments are detailed in the LBA section.	
6	It was supposed that new European guideline will follow a trend to harmonize and improve 9 years old FDA Guideline. Although most general requirements have been formally followed the differences between the FDA and EMEA for each validation tests are significant. Hence, it may cause problems to adopt general approach meeting both FDA and EMEA requirements. In our opinion, the proposed guideline does not improve quality of the analytical method compared to methods validated according to FDA although the total number of experiments is significantly increased that, consequently, will raise dramatically the total analytical cost.	The EMA considers harmonisation an important topic and has therefore embarked into numerous discussions in relation to the preparation of this guideline. The outcome of these interactions on harmonisation issues have been taken into account as much as possible.
8	There is no mention of recovery as being a requirement for validation. Does this mean that recovery will no longer need to be validated?	Extraction recovery is an issue investigated during the analytical method development and as such considered not necessary to be included in this guideline.
9	The guidelines should be further refined to provide more specific guidance for the validation of ligand binding assays (LBA). Many of the comments are specific for chromatographic assays, but the section on ligand-binding assays indicates that same validation principles generally apply. However, given the wide application of ligand binding assays, we suggest that the same level of detail should be provided in the guidance for both chromatographic and ligand binding assays. We therefore suggest amending the individual sections in the guidance to provide the same level of specific detail for both chromatographic and ligand binding assays, indicating where items apply to both types, or only one, assay modality, and stating alternative specifications for ligand binding assays when required.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
11	1). A significant part of this guideline deals with the analysis of study samples. Therefore, this should be reflected in the title of this guideline. 2). There were mixed usage of "bioanalytical methods" and "analytical methods" in this guideline. It is better to be more consistent to use	1. The title of the guideline has been slightly adapted, however the title is kept short. No separation has been made between the different parts of this guideline, as it is considered that the title is clear and reflects the contents of the guideline.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	only one of them and define them early in the guideline.	2 This is acknowledged and wording should be consistent. The wording will be changed, if applicable.
15	Baxter BioScience develops and manufactures therapeutically used plasma-derived and recombinant macromolecules including coagulation factors, immunoglobulins, protease inhibitors and sealant proteins as well as vaccines. Preclinical and clinical research therefore focuses on macromolecules. The following comments on the Draft Guideline on Validation of Bioanalytical methods, issued Nov 19, 2009, were compiled by a team of scientists with expertise in preclinical and clinical research required for the characterization of macromolecules, the analytical testing of samples derived from such studies and in QC release testing of plasma- and recombinant products. The analytical methods covered by our experience are coagulation assays, enzyme activity assays, ligand-binding assays, immunogenicity testing but not LC-MS methods as they are not commonly used forPK and TK measurement of macromolecules. It was noted that this guideline is apparently designed for • R&D purposes but not for quality control purposes related to the drug and that we would ask for a clearer statement to confirm this. • Chemical, including immunological-, ligand binding-, and similar methods but not for biological methods such as coagulation assays and similar methods as decribed in Chapter 2.7 of European Pharmacopoeia. A clear statement confirming this is also missing. It should also be pointed out that a guideline for biological methods (as defined above), that also covers the evaluation of these assays as described in chapter 5.3 of European Pharmacopoeia is still missing and certainly remains a requirement. A Working Group to establish a separate guideline for this scope would be recommended and Baxter would offer to provide experts as participants in this working group. We would also like to call attention to the ongoing initiatives in the US, where FDA and AAPS agreed on a consensus document on bioanalytical method validation during the 3rd FDA/AAPS workshop. Our team identified the following major points for discussion:	The section on the scope of the guideline has been changed to better reflect the scope of this guideline. Methods used for determining quantitative concentrations of biomarkers used in assessing pharmacodynamic endpoints are out of the scope of this guideline. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
16	It is proposed that requirements for analysis of study samples from bioequivalence trials and from preclinical safety studies (GLP required by legislation) should be differentiated from requirements for the	- For GLP issues, we refer to the comments given on the specific questions addressing this issue.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	analysis of samples from other trials (e.g. animal PK studies, population PK studies, clinical pharmacology and clinical safety and efficacy studies). An argument in support of this proposal is that for approval of a generic drug which is only based on one pharmacokinetic bioequivalence study, a plausibility check of the bioanalytical results considering the pharmacokinetic outcome is only within limits possible. However, approval of a new drug requires very often 50 and more studies with a pharmacokinetic component allowing a profound cross-check of the bioanalytical results. Coherent wordings over the whole guideline text: e.g. "full" validation instead of "complete" validation will facilitate understanding of requirements.	- This is acknowledged and wording should be consistent. The wording was changed as appropriate.
17	A more clear distinction of the requirements for toxicokinetic studies, clinical studies and especially BE studies should be given.	We are not aware that there is a distinction in requirements.
18	The Guideline is generally consistent with international practices and the CHMP is encouraged to maintain that alignment during finalization.	Comment noted.
19	This guideline defines several thresholds and deviations (e.g. "20% of...."). Proposal: It should be pointed out that variations of these values are possible, if they are properly justified by data obtained from assay development.	Analytical methods should fit for the purpose the methods are used for. Acceptance criteria wider than those defined in this guideline may be used in special situations. This should be prospectively justified based on the intended use of the method.
20	It was supposed that new European guideline will follow a trend to harmonize and improve 9 years old FDA Guideline. Although most general requirements have been formally followed the differences between the FDA and EMEA for each validation tests are significant. Hence, it may cause problems to adopt general approach meeting both FDA and EMEA requirements. In our opinion, the proposed guideline does not improve quality of the analytical method compared to methods validated according to FDA although the total number of experiments is increased.	The EMA considers harmonisation an important topic and has therefore embarked into numerous discussions in relation to the preparation of this guideline. The outcome of these interactions on harmonisation issues have been taken into account as much as possible.
21	EFPIA welcome the European Medicines Agency initiative to release a guideline on validation of bioanalytical methods at the time when an update of the 2001 FDA guidance on the same subject is expected.	The EMA considers harmonisation an important topic and has therefore embarked into numerous discussions in relation to the preparation of this guideline. The outcome of these interactions on harmonisation issues have been taken into

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	However, we strongly advocate that a single approach to bioanalytical methods validation and sample analysis is agreed upon globally, in order to allow that a single standard is followed by companies submitting global marketing applications We would like to avoid a situation where FDA, EMA and other Regulatory Agencies provide divergent guidance that will make method validation and routine analysis a task even more complex than it is because of its nature. In order to assist bioanalytical laboratories to fully appreciate the sections that are relevant to their area, it is recommended that the large molecule (ligand binding assays) and small molecule (chromatography assays) be clearly separated in this document. While it is appreciated that there is a section specifically on Ligand Binding assays, it is not clear for the items not covered in this section if the rest of the document is therefore applicable to Ligand Binding assays. This needs to be reflected more in the guidance at several places and we would advise to make absolutely clear which recommendations hold for which molecule class or analytical technique. We also recommend that a clearer distinction is made between the requirements pertaining to validation and routine sample analysis. In the current draft it is often difficult to distinguish between the two and we would recommend that clearly distinguished sub-paragraphs are introduced for the sake of clarity. The scope of the draft Guideline is very broad as it covers both analytical method validation and sample analysis in biological matrices, and that the analyte could be anything from an organic synthetic molecule, a protein and a nucleic acid to a living microorganism, that the application of the analyses mentioned spans over a wide range where toxicokinetics, pharmacokinetics and bioequivalence and that both GLP and GCP are mentioned as referential, with the reporting according to GLP being further detailed. We would recommend that either the scope of the Guideline be better defined and limited or that the above-mentioned areas and topics be	account as much as possible. - It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. The title of the guideline has been slightly adapted, however the title is kept short. No separation has been made between the different parts of this guideline, as it is considered that the title is clear and reflects the contents of the guideline.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	fully covered in relation to the OECD GLP requirements and the Agency expectation in terms of data to be generated and presented in Marketing Authorisation Applications. Section 7 attempts to address several types of reports and it is suggested that each of the following reports have their own section: Study Reports, Assay Validation Reports, Bio-Analytical Reports. In addition, this section seems to have a more non-clinical focus and the Agency may want to also consider how to the recommendation could apply to clinical studies or clearly specify that the recommendation in this section only apply to non-clinical related analyses.	
22	The European Generic medicines Association (EGA) welcomes the European Medicines Agency (EMA) draft guideline on bioanalytical method validation. Although in general we support the approaches laid out in this document, we believe certain areas remain to be clarified in order to achieve an optimal implementation. The relevant points are discussed in the next section 'Specific comments on text'. The EGA is supportive of the idea of having one unique general guideline covering all aspects of bioanalytical method validation regardless of the type of measurement of drug concentrations in biological matrices with the possibility to justify deviations case by case. Other comments: The EGA further identified two topics which we would recommend to include in the final guideline given their importance: they are used today by EU medicines agencies to support critical deviations during inspections: The proposed section to be included relates to Internal standard variability and sample re-injection. 1 - Internal standard variability The response of the Internal Standard (IS) should be controlled in each batch. Study samples that do not comply with the criteria	1. - It is difficult to give criteria on IS response and deviations. This may be method dependent.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	indicated in the SOP of each laboratory should be re-analysed. As a general rule, the response of the IS in the study samples should be within ± 50% of the average IS response obtained in the accepted calibration standards and quality controls. However, for analytical methods with high variability (i.e. including a derivatization step), it should be acceptable to omit the acceptance criteria of the IS response described above, provided the variability of the analytical method is corrected by using a stable isotope labelled internal standard. In this case, the following study samples shall be rejected and re-analyzed: - Samples presenting a response of the IS lower than 5% of the average IS response obtained in the accepted calibration standards and quality controls. - Samples where the analyte presents a signal/noise ratio (S/N) < 5. A cross-validation with QC and real samples is a definite tool to validate a method when there may be doubts about its reliability in a specific laboratory. It is a conclusive test to determine the adequate reproducibility of the results. The EGA would welcome these principles to appear in the final text of the guidance document. 2 - Sample re-injection Provided samples have to be re-injected due to poor chromatography, it is not clear whether quantification should be carried out using calibrators of the original batch or using calibrators of the sequence with which the injection is performed. In large studies (e.g. with thousands of samples and/or long chromatography times), in the event the analytical run has to be stopped at some point to solve small chromatographic system problems (for example those related to connections, valves, etc.), it is not clear whether the entire batch has to be re-injected or whether the injection procedure should be re-started at the moment when it was stopped (evaluating the samples with controls distributed before and after the stop).	Analytical methods should fit for the purpose the methods are used for. Acceptance criteria wider than those defined in this guideline may be used in special situations. This should be prospectively justified based on the intended use of the method. - Criteria set for routine analysis of study samples should be in line with those applied/obtained during validation. 2. No comment is given, as it is difficult to give comments on specific analytical issues, as they may be case dependent.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	Considering the above depicted situations, when there are issues of poor chromatography identified in a set of samples of an analytical batch, the EGA would recommend to include the provision that sample re-injection should be allowed along with that of the quality controls showing the same chromatography problems as the set of affected samples.	
23	Comment: This guideline does not only address aspects concerning validation, but also of routine analytical run acceptance see section 5. The title of the document is therefore not totally adapted to the content. **Proposed change (if any):** Change to e.g. "Guideline on validation and acceptance criteria for bio-analytical methods and the analysis of study samples"	The title of the guideline has been slightly adapted, however, the title is kept short. No separation has been made between the different parts of this guideline, as it is considered that the title is clear and reflects the contents of the guideline.
	Comment: Method validation is different between physico-chemical analysis and bio-assays. We therefore suggest to have either a specific guideline for bio-assays or a specific paragraph no sensitivity, carry-over, LOQ, calibration curves, accuracy, precision, dilution integrity, matrix effects and stability.	The comment is unclear. However, a separate section has been introduced on ligand binding assays.
	Comment: The present scope of the Guideline is huge due to the fact that: - it covers both analytical method validation and sample analysis in biological matrices - the analyte could be anything from an organic synthetic molecule over a protein and a nucleic acid to a living microorganism - the application of the analyses mentioned spans over a wide range where toxicokinetcs, pharmacokinetics and bioequivalence are mentioned - both GLP and GCP are mentioned as legal authority referentials, and the reporting according to GLP routines is further detailed, although GLP compliance cannot be claimed.	The comment is noted.
	Comment: It is felt that the Guideline needs a thorough review and update either: a) to fully cover this complex field and more clearly describe what the EMA expects in the regulatory applications from industry in relation to the existing OECD GLP requirements	

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	b) or to introduce more clear limitations concerning the scope of the Guideline. Comment: While looking through the guideline it is also noticed that some terms are not consistently used throughout the document. To avoid confusion we recommend to harmonize within the document and to include, as necessary the definition of terms. This concerns, for example: • "protocol" vs. "plan" (line 388 "study plan"; line 397/398 "analytical study protocol" vs. "study plan", line 439 "study protocol", line 491 "protocol", line 496 "analytical protocol") • "in compliance" vs. "in accordance" (line 69/70, 489, 492) What do you mean with these terms? Please define it inside this guideline for clarification.	This is acknowledged and wording should be consistent. The wording will be changed, if applicable.
24	The guideline appears to be strongly focusing on small molecules (chromatography based assays, e.g., LC-MS) bioanalysis. Large molecules are poorly represented so we assume this guidance is applicable for large molecules too. This needs to be reflected more in the guidance in several places. We ask for more clarity on which recommendation hold for which molecule class or analytical technique. Examples where this is most apparent: acceptance criteria of 15%CV vs. 20% CV, metabolites and degradation products. We would like to ask EMA to use either an integrated approach (chromatographic and ligand-binding methods) or separate chapters for each technique of the same discipline. This guideline does not only address aspects concerning validation, see section 5. The title of the document is therefore not totally adapted to the content. We would propose to change to" Guideline on bioanalytical procedures: assay validation and routine analysis" The guidance is addressing method validation and study conduct using chromatography based (e.g., LC-UV, LC-MS, GC-MS,…) and ligand-binding based (LBA) bioanalytical assays. This may be reflected in the title and as well in the chapters where both techniques are addressed and only different acceptance criteria are applicable (e.g., 15 vs 20%). Method specific recommendations may be addressed in special chapters. We would like to ask EMA to change the wording for samples used during the validation from QC to validation or stability samples, since	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. The document has been checked for consistency and wording

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	they could be identical but doesn't need to be necessarily. We would like to consider EMA to use terms consistently throughout the document: "Quantification" / "Quantify" instead of "Quantitation" / "Quantitate"; "Full validation" instead of "complete validation" "LLOQ" vs "LOQ" "Actual analysis" vs "routine analysis" "Inter-assay" and "intra-assay" vs "between run" and "within run" (precision and accuracy) The document did not mention duplicate analysis (incubation of samples) in LBAs, this topic maybe addressed. We would like to ask EMA to provide a glossary of term which is consistent with FDA	has been changed, if applicable.
26	Important topics addressed in the guideline are generally similar to those in the 2001 FDA guidance. This guideline includes a discussion of incurred sample reanalysis which is an important addition given discussions and regulatory input on this topic in the last few years, although some more specific guidance in this area could be provided (see specific comments). The general layout of the guideline is logical and easy to follow and represents a step in the process of international regulatory harmonization on this topic. Perhaps more emphasis should be given to some of the differences between chromatography-based and ligand-binding assays (LBAs). The LBA community will want to point out some of these differences [such as detection being indirect (LBAs) versus direct, the heterogeneity of reference standards (LBAs), typical absence of extraction prior to analysis (LBAs), etc.], and how these differences should perhaps lead to either separate documents for the two types of assays or a clearer demarcation within one guideline.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
30	You may consider it useful to estimate the total measurement uncertainty in addition to determine CV% values, in order to understand measurement uncertainty and corresponding tolerance intervals under worse case conditions. This is especially necessary when comparing measurements to limit specifications.	The comment is unclear.
32	Outline / Table of contents In general our specific comments reflect some of those presented at	

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	the EBF-EUFEPS meeting with some modifications. In addition we feel it is important to ensure the wording /definitions are aligned with current usage and common practice such as ; Full validation vs complete validation Actual analysis vs study sample analysis There should be a clear distinction between data that needs to be included in the study file from that mentioned in the study report. Within the Guideline there needs to be definitive separation of the sections that relate to development, validation and sample analysis. Where appropriate, requirements specific to bioequivalence/bioavailability studies should be highlighted. We agree with the comments made at the BF-EUFEPS meeting in Brussels:"It would be suggested that the section that discusses the reference standards should be included into a separate chapter since this is not specifically related to validations."	 The document has been checked for consistency and wording has been changed, if applicable. Study file requirements are issues for GLP. Analytical method development is outside the scope of this guideline. Bioequivalence specific issues will be identified in this guideline, if applicable. The paragraph on Reference standards has been kept under section Method validation, as it is considered that it is clear that this refers to validation and other analytical issues.
34	Bioanalytical methods is a word that lacks clarity. It is not clear if this guideline provides requirements for the validation of biomarkers (example: cytokines release), assays to assess immunogenicity of the drug product or for example consequence of a drug treatment (number of residual cancerous cells after monoclonal antibody therapy). The scope of this guideline must be much more detailed.	The title of the guideline has been slightly adapted, however the title is kept short. No separation has been made between the different parts of this guideline, as it is considered that the title is clear and reflects the contents of the guideline.
36	Pfizer welcomes the opportunity to provide EMEA with comments on this draft guidance. Pfizer hopes that the EMEA guideline will generally match as close as possible the Guidance for Industry Bioanalytical Method Validation from the U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Veterinary Medicine (CVM), published in May 2001 (http://www.fda.gov/cder/guidance/4252fnl.pdf) supplemented with the outcome of 2006 Workshop/Conference - Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays	The comment is noted.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	(http://www.aapsj.org/view.asp?art=aapsj0901004), and the references cited therein. Additionally, regarding Ligand Binding Assays (LBAs), We would also proposes that the EMEA guideline is harmonized with the additional recommendations given in DeSilva B et al. Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. Pharm Res. 2003; 20: 1885-1900). nes or recommendations.	
37	In our opinion, the reliable results of bioanalytical part of the study are crucial for the proper decision on drug registration or rejection. Therefore EMA guideline on this topic seems to be necessary. As stated in introduction, the quality of measurement results for a specific application should be the ultimate goal of the guideline. In our opinion, to prevent possible misunderstanding between regulatory authorities and pharmaceutical industry, the guideline should rather define requirements than provide recommendations. It should be clearly stated what is required and that another approach than recommended is acceptable. We propose to transfer recommendations to different section of guidance then requirements or for example to write recommendations in italics. Currently, as there were no detailed EMA guideline on bioanalytical method validation, the bioanalytical laboratories are basing on FDA document "Guidance on bioanalytical method validation" issued in May 2001. Taking into account global perspective, represented by ICH, we strongly suggest that EMA guideline should be as much harmonized with FDA requirements as possible. The differences, which are in some points necessary, should have strong scientific background. The final guideline will definitely influence possibility of drug registration in EU and USA basing on the same bioanalytical method. In case FDA acceptable bioanalytical method validation will be unacceptable for EMA, the repeated clinical trial will be necessary. The ethical responsibility of unnecessary studies in vivo, especially in humans, will be definitely on EMA site. The additional and unnecessary cost of extra trials will be shifted by pharmaceutical companies to the customers leading to further increase in public health system spending.	The EMA considers harmonisation an important topic and has therefore embarked into numerous discussions in relation to the preparation of this guideline. The outcome of these interactions on harmonisation issues have been taken into account as much as possible. The comment is noted.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	In our opinion, more risk-based approach should be used in case of bioanalytical methods validation and the final scope of the study should be taken into consideration. For example bioequivalence studies are definitely comparative studies, so the request for the analysis of incurred samples may be questionable: method should allow comparison of two data sets (i.e. bioanalytical results) not determination of absolute concentration. The constant ratio of Cmax between test and reference drug is of primary interest rather than identical concentration measured during 2 separate analysis of each sample.	It is considered that a risk based approach is not applicable. Analytical methods should provide reliable and accurate information. This is confirmed by additional incurred sample reanalysis. A risk based approach can be applied with the results of the study and the outcome of the study itself.
	The required content of printed study reports should be as limited as possible. We suggest that for example representative chromatograms should be attached to report in electronic (PDF) version only. In case any questions arise during documentation assessment by regulatory agency, further detailed data in electronic format should be available at request.	See comment on questions regarding this issue.
	Title "Guideline on validation of bioanalytical methods" does not correspond to the nomenclature used inside, where term "analytical method" rather than "bioanalytical method" is used. We recommend constant use of term "bioanalytical method".	The document has been checked for consistency and wording has been changed, if applicable.
	The information on LOD (limit of detection) and recovery are missing. If they are not necessary for registration purposes it should be clearly stated in the guideline. In our opinion both parameters are useful during method development, but both are not necessary to ensure efficacy and safety of studied drugs.	Limit of detection and recovery are issues investigated during the analytical method development and as such considered not necessary to be included in this guideline.
	In some cases, especially precision, accuracy and stability mathematical formulas should be presented, which will definitely lead to better understanding of definition of terms.	See comments on questions regarding this issue.
	To facilitate the interpretation of data by regulatory agencies as well as to improve data quality, especially in case of stability studies, wider application of confidence intervals is advised [U. Timm, M. Wall, D. Dell, J. Pharm. Sci. 74 (1985) 972-977; P. Rudzki, A. Leś, Acta Pol. Pharm. 65 (2008), 743-747]. Confidence intervals for mean accuracy can replace accuracy and precision determination as two separate	Precision/accuracy calculation methods with confidence. intervals may be used for explorative issues, but considered not a criteria to be applied for validation.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	points of validation.	
39	We do not see a need for this new guideline. Of course, the bioanalytical methods must be validated according to the current scientific knowledge, but with every new guideline the scientific approach is more and more replaced with a formal approach. The quality of drug products will not be improved with this new guideline, the current validation processes of bioanalytical methods are satisfactory and general consensus on basic validation principles exists in literature. If problems with validation arise, the inspectors can reject the bad method based on scientific arguments. The new guideline will only burden the analytical departments with new formal requirements without improving quality of the methods. We enclose our comments to the guideline, but we recommend to stop this project and keep the current status.	The EMA considers harmonisation an important topic and has therefore embarked into numerous discussions in relation to the preparation of this guideline. The outcome of these interactions on harmonisation issues have been taken into account as much as possible.
45	We thank the European Medicines Agency for putting together this document. We think that this document is very important, and a timely update of the 2001 FDA guidance on the same subject, showing the current status of the art in this area. We strongly advocate that a **single** guidance is agreed upon by the major Regulatory Agencies of the most important countries, in order to allow that a single standard is followed by companies submitting global marketing applications We would like to avoid a situation where FDA, EMEA and other Regulatory Agencies provide divergent guidance that will make method validation and routine analysis a task even more complex than it is because of its nature. We strongly support harmonization of the regulatory effort between the two agencies and with existing AAPS "White Papers". Recommendation: please consider a "harmonized" global approach to Guidances on Bioanalytical Method Validation In order to assist bioanalytical labs to fully appreciate the sections that are relevant to their area, it is recommended that the large molecule	The EMA considers harmonisation an important topic and has therefore embarked into numerous discussions in relation to the preparation of this guideline. The outcome of these interactions on harmonisation issues have been taken into account as much as possible. It is recognised that chromatographic methods and ligand

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	(ligand binding assays) and small molecule (chromatography assays) be clearly separated in this document. While it is appreciated that there is a section specifically on Ligand Binding assays, it is not clear for the items not covered in this section if the rest of the document is therefore applicable to Ligand Binding assays. In addition, it should be noted that a validation plan with acceptance criteria should be drafted and approved prior to the execution of the validation of the assay. Section 7 attempts to address several types of reports and it is suggested that each of the following reports have its own section: Study Reports, Assay Validation Reports, BioAnalytical Reports. In addition, this section seems to have a more non-clinical focus and the agency may want to also consider how to consider for clinical studies as well; unless this indeed ONLY applies to non-clinical related. Suggest that Incurred Sample Reanalysis studies be a stand alone appendix to study reports and not be part of the Bioanalytical Report. There is often a mix of validation and routine analysis requirements in the paragraphs. This should be separated more clearly	binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. See comments on questions regarding this issue.
47	We have provided a description of comments with recommendations specific to lines and sections in the draft guidance, in the table following. In the table, we identify the line number of the draft guidance; the middle column a respective comment and rationale, and the third column, a suggested change. We would be pleased to provide additional clarification of comments, upon request. Additionally, we have provided in this cover letter a summary of what we feel are three significant perspectives that we hope are considered when revising the draft document. Although the response is supported with scientific justifications for its positions, comments and recommendations, the flavor of this reply is that of a quality perspective with a foundation of auditing experience both within the United States, and more importantly, globally. The SQA and BASS thank the CHMP and EMEA for generating a guideline addressing bioanalytical method validation, with emphasis on LC-MS/MS and immunoassays. Such a document is long overdue	Comment noted. The EMA considers harmonisation an important topic and has therefore embarked into numerous discussions in relation to the preparation of this guideline. The outcome of these

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	since outside of the FDA guidance and the Crystal City III White Paper, there is no other regulatory advice available, which is significant now as few medicines are developed for submission solely in the USA. We also appreciate that many representatives of the EMEA and CHMP are supportive of a harmonization, and even globalization of bioanalytical method validation (BMV) guidance, and respect this initiative in the preparation of the applicable draft guidance. Since the SQA shares the same position regarding globalization, the following comments are given with this consideration, and that of the FDA's intention of updating its own BMV guidance document in 2011. We suspect that the content of the EMEA guidance can impact the FDA document, and the rate at which the harmonization initiative can proceed. 1. Insufficient treatment of the ligand-bindings assays. While the scope of this guidance document is the validation of bioanalytical methods, ligand-binding assays (LBA) are not adequately treated throughout this document, although it is apparent that is the intention of the EMEA. All method types and associated acceptance criteria for LBA should be addressed in this guidance. For instance, the method parameters described throughout Section 4.1 (Complete validation of an analytical method) fully address parameters for LC/MS and provides only limited guidance for immunoassays/ligand-binding assays, and the Ligand-Binding section lacks significantly in its content pertaining to ligand-binding assays where routine requirements such as cut points, dilution linearity/parallelism, the correct use of non-linearity and anchor points are not described, to name a few. The context of the document should treat LBA and LC/MS acceptance criteria similarly, even if there are circumstances with LBA methods that are more variable and thus allow for greater flexibility in acceptance parameters. This is already a best practice in industry, and thus the guidance should reflect this, especially since the technologies associated with immunoassays are advancing as rapidly as the complexity of the associated macromolecules. Furthermore, it is not clear as to whether the EMEA draft guidance is supposed to address immunoassays, which could also implicate ADA/immunogenicity, receptor assays, biomarker assays, etc., and this needs to be clarified within this document, and similarly be consistent.	interactions on harmonisation issues have been taken into account as much as possible. 1. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	Many in the industry conducting LBA and macromolecular analyses believe that immunoassays and LBAs are being treated like a footnote or add-on in guidance documents on analytical method validation. There have been suggestions about removing this section altogether, which we would not object to. Unfortunately, this would leave a void within the bioanalytical method validation paradigm associated with GLP and human clinical bioanalysis. The period in which a guidance could otherwise be prepared to address LBA and other immunoassays is not predictable, but likely too long. Therefore, there is an immediate need for a guidance document that would extend beyond a stated understanding that the basic principles of validating bioanalytical methods for small molecule also apply to ligand-binding assays except for a variation in the types of validation method parameters that require investigation and documentation, and the allowance for looser acceptance criteria because of the unique heterogeneity and other aspects of the immunoassay (supported with scientific rationale). We feel that the EMEA BMV guidance offers the ideal opportunity to provide guidance relevant to ligand-binding assays, and even other immunoassays including ADA and biomarker methods. Having such a guidance would facilitate consistent regulatory approaches to ensuring an adequate method, such as scientific peer review and QAU/agency auditing, while allowing for the creative science necessary to facilitate effective drug development and appropriate drug submissions. It is our recommendation that a similar approach and language to immunoassays (at least LBA) is used throughout the EMEA BMV guidance as currently described for LC/MS methods. 2. Need for the definition of validation and qualification. The requirements for varying degrees of validation and qualification are becoming necessary with the current and global development of more diverse types of medicines. Well defined definitions for validation and qualification would add significant clarity to this guidance. Not only is the global industry currently able to develop medicines associated with non-small molecules such as proteins, conjugated antibodies-drugs, iRNA, and peptide mimetics to name a few, the technologies and bioanalytical paradigms associated with	2. Analytical methods should fit for the purpose the methods are used for. Acceptance criteria wider than those defined in this guideline may be used in special situations. This should be prospectively justified based on the intended use of the method. Methods used for determining quantitative concentrations of biomarkers used in assessing pharmacodynamic endpoints are out of the scope of this guideline.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	these have advanced significantly in recent years. As far as regulating the methods associated with these new paradigms is concerned, it is difficult to discern whether a method should be validated, or qualified, and if a method can be qualified, to what degree? For instance, a bioanalytical method developed for the quantitation of drug in urine that is not controlled for collection conditions cannot be validated, but qualified, and likely qualified to a greater degree than a biomarker assay. Interestingly, biomarker assays are now becoming quantitative, utilizing various mass spectrometry techniques, and thus can be validated. Furthermore, auditors in the industry continue to confront facilities that use commercial kits that come 'qualified' for an intended use, but nonetheless use the said kits as a 'validated' method. The validation of a kit needs to be defined if the EMEA wishes to distinguish this from a qualified method. If a commercial kit is not validated to begin with, it cannot be re-validated, as the current draft suggests. A kit is used for its materials/reagents during a method validation, and this should be clarified. Since the science associated with these paradigms are advancing faster than consensus papers can address them from a quality and regulatory perspective, addressing the practices and applications of validation and qualification would be very important to this guidance document. 3. Lack of clarity in regulatory applicability to bioanalysis of samples from clinical studies. If the scope of the guidance is intended to include human clinical samples, then the guidance should be precise in this aspect of application and provide the respective details. Within the attached comments and recommendations, we recommended that the guidance indicate that the guidance described is applicable to human clinical samples as they apply to bioanalytical method validation, and further suggest making a distinction between GCP and GLP requirements in the guidance. This is a very important issue as far as SQA and BASS are concerned as there is currently little guidance regarding the bioanalysis of human clinical samples. Since a BMV guidance is to facilitate recommendations and/or requirements to assure accuracy and sufficient sensitivity to measure, with appropriate precision, the actual concentration of the active drug ingredients, therapeutic moieties, and/or its active metabolite(s) in a biological sample, the	3. Comment noted.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	application of the principles described in an EMEA BMV document should be similarly applicable to human clinical samples. Recognizing the difficulties concerning the regulatory status of analysis of human clinical samples, the British Association of Research Quality Assurance (BARQA) published a document on 'Good Clinical Laboratory Practice' in 2003, which described a framework for the management and documentation of clinical bioanalysis, based on GLP principles. This document followed one forwarded by the EMEA Committee for Proprietary Medicinal Products (CPMP) adopting a revised Note for Guidance on bioavailability and bioequivalence that came into effect in 2002. This guidance stated that the bioanalytical part of bioequivalence trials should be conducted according to the applicable principles of GLP. The MHRA has offered some additional guidance in 2009 with their GCP "Guidance on the maintenance of regulatory compliance in laboratories that perform the analysis or evaluation of clinical trial samples." Therefore, there is already some framework for clarification regarding guidance for the bioanalysis of human clinical samples, and thus making reference to GLPs within the guidance, only when and if it is applicable. It is the hope of the SQA and BASS that the EMEA guidance paves the way for this needed clarity.	
51	QC preparation CV for curve points Formulae	The comment is unclear.
52	PPTA member companies develop and manufacture plasma-derived and recombinant macromolecules including coagulation factors, immunoglobulins, protease inhibitors and sealant proteins. Preclinical and clinical research therefore focuses on macromolecules. PPTA's experts noted that this guideline is apparently designed for • R&D purposes but not for quality control purposes related to the medicinal product. We would appreciate a clearer statement to confirm this. • Chemical, including immunological-, ligand binding-, and similar methods but not for biological methods such as coagulation assays and similar methods as described in Chapter 2.7 of European Pharmacopoeia. A clear statement confirming this is also missing. It should also be pointed out that a guideline for biological methods (as defined above), that also covers the evaluation of these assays as described in chapter 5.3 of European Pharmacopoeia is still missing	The scope of the guideline has been changed. Methods used for determining quantitative concentrations of biomarkers used in assessing pharmacodynamic endpoints are out of the scope of this guideline. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	and certainly remains a requirement. A Working Group to establish a separate guideline for this scope would be recommended and PPTA would offer to provide experts as participants in this working group. We would also like to call attention to the ongoing initiatives in the US, where FDA and AAPS agreed on a consensus document on bioanalytical method validation during the 3rd FDA/AAPS workshop.	

2. Specific comments on text

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
7	21	The title of the draft guideline only mentions method validation. However more topics are covered in the document including sample analysis. **Proposed change (if any):** We would suggest revising the title of the guideline to reflect its actual content once finalised, e.g. "Guideline on validation of bioanalytical and sample analysis".	The title of the guideline has been slightly adapted, however the title is kept short. No separation has been made between the different parts of this guideline, as it is considered that the title is clear and reflects the contents of the guideline.
7	41	Comment: Only guideline for validation, but the application to bioanalytical studies is also included **Proposed change (if any):** title to be changed to 'guideline on validation of bioanalytical methods and application to routine use'	The title of the guideline has been slightly adapted, however the title is kept short. No separation has been made between the different parts of this guideline, as it is considered that the title is clear and reflects the contents of the guideline.
39	36	**Executive summary** Pfizer suggests that greater clarity is introduced in the document to define where particular analysis platforms have differing requirements. In particular, the document should discriminate where the requirements and discussions are not relevant to a technique such as Ligand binding assays (LBA). For example, statements in the sections on assay selectivity, carryover, matrix effects are relevant for the LC-MS/Ms platform only Greater clarity is sought on the scope of the guideline as to whether quantitative biomarker data is in the suggested scope of the guideline and if any differentiation is sought between "small", "large" or biotherapeutic molecules or entities. More clarity should be considered in better defining the scope of what activities are considered validation and what activities are method development and thus are	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. Methods used for determining quantitative concentrations of biomarkers used in assessing pharmacodynamic endpoints are out of the scope of this guideline. This is clearly indicated now. All issues in this guideline are related to validation and should be reported. Analytical method development is outside the scope of this guideline. The section on the scope of the guideline has been changed to better reflect the scope of this guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		not required as reportable within a validation report. Additionally, as specific aspects of the application of the bionalytical methods in the analysis of samples in studies and clinical trials are addressed, it is suggested text is added to the executive statement on these requirements. **Proposed change (if any):** Suggest rewording "the guideline focuses on the validation of the analytical methods used for pharmacokinetic sample analysis" to " The guideline focuses on the validation of analytical methods used in generating quantitative concentration data used for pharmacokinetic and toxicokinetic parameter determinations. Guidance is given on the application of these validated methods in the analysis of samples from animal and human studies and trials. Methods used for determining quantitative concentrations of biomarkers used in assessing pharmacodymanic endpoints are out of scope of this guideline.	
39-43	2	• Comment: This guidance is primarily for chromatographic-based methods. It would be helpful if the applicability of this guidance to ligand-binding assays be addressed up front in the document. • The acceptance criteria between the LC-MS and LBA are different but the opening summary does not clarify this. • Guidance is strongly focusing on small molecules (LC-MS) bioanalysis. Large molecules are not excluded so we assume this guidance is applicable for large molecules too. This needs to be reflected more in the guidance at several places. It is advised	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. Including general applicable parameters in this section is considered not the correct place and too detailed for this section.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		to make more clear which recommendations hold for which molecule class or analytical technique. Examples where this is most apparent: acceptance criteria of 15%CV vs. 20% CV, metabolites and degradation products. **Proposed change (if any):** • It should be clearly stated what are general applicable parameters for all bioanalytical methods and then consider having separate sections specific to chromatographic methods (e.g LC-MS) and ligand binding methods (e.g. ELISA).	
39-43	21	The draft guideline defines key elements and provides recommendations for the validation of bioanalytical methods. We would suggest that a differentiation between method development and validation should be made. Not all items need to be addressed as part of method validation, some could be addressed during method development and do not need to reported in a method validation report.	Analytical method development is outside the scope of this guideline. All recommendations included in this guideline should be addressed as validation requirements.
39-43	23	Executive summary do not cover whole content. **Proposed change:** Adjust Executive Summary to content of Guideline e.g. the guideline is about bioanalysis and it is more appropriate to use "bioanalytical" instead of "pharmacokinetic" in line 41. Please delete sentence "The guideline focuses on….." since we at present do not have different validation guidelines for a method depending on the purpose of the study.	The executive summary has been changed to better reflect the content of the guideline.
39-43	26	Executive Summary: This section appears to say little and does not really provide a summary of the document contents. **Proposed change (if any):** Either expand the	The Executive summary section is a part of all CHMP guidelines. The section is not expanded, as this may lead to repetitions.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		summary to be meaningful or delete it	
39-43	45	**Proposed change (if any):** Replace "actual analysis" by "routine analysis"	The section has been changed including 'routine analysis'.
40-41	37	Comment: see general comments **Proposed change (if any):** This guideline defines key elements and describes requirements for the validation of bioanalytical methods.	The text has been changed into**: '**This guideline defines key elements necessary for the validation of bioanalytical methods.'
44	36	**Introduction** The document indicates that wider acceptance criteria may be used "in special circumstances such as analysis of complex matrices" The proposed guideline does not give sufficient detail on the agencies thinking as to what constitutes complex matrices. Furthermore the there is no discussion as to how this point nor does it sufficiently address the "tired approach" for metabolite quantification. There is no discussion in the introduction of the need for failed run investigations and how those are to be included in method validation and method application.	The introduction section has been changed. The example of solid tissues has been deleted to prevent misunderstanding. Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be subject to case by case evaluation. Failed runs discussion is not a topic to be included in the introduction section.
44 -51	24	Comment: We would like EMA to consider the introduction of the "tiered approach" for biomarker and early metabolite quantification as well as quantification in alternative matrices (e.g., tissue) where rules for a full validation are not yet applicable. We would prefer a much clearer separation between method development, method validation and routine analysis.	Methods used for determining quantitative concentrations of biomarkers used in assessing pharmacodynamic endpoints are outside the scope of this guideline. This is clearly indicated now. Early studies are not covered by this guideline. As indicated, we are referring to toxicokinetic studies. Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be subject to case by case evaluation.
44-54	2	Comment:	The introduction section has been changed. The example of

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Lines 52-54 suggest inclusion of validation for tissue analysis within the validation guidelines. It is normally not possible to measure extraction efficiency or determine accuracy or stability in solid tissues. One can determine such parameters only in tissue homogenates. Also, the collection of solid tissues is not a controlled, well defined process, and tissues are heterogeneous. The solid tissue concentration measurement should, therefore, not require regulated bioanalysis, but concentrations could be measured for research information purpose only.	solid tissues has been deleted to prevent misunderstanding.
44-54	7	Comment: In certain situations, the use of qualified methods should be allowed, i.e. with respect to metabolite quantification and "special" matrices such as tissues. The tiered approach is supported by the other regulatory agencies, such as the US FDA.	Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be subject to case by case evaluation.
44-54	21	Comments: The document does not address a "tiered approach" for metabolite or biomarker quantification or for alternative matrices such as tissues and urine as supported by FDA. Additionally, the draft guideline does not address the issue of Failed Run Investigations. We understand that the revised FDA Guidance will very likely cover these and would suggest that every effort be made to harmonize the new CHMP guideline in this respect.	Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be subject to case by case evaluation.
44-62	23	Comment: This guideline only applies to the Human Pharmaceutical Industry. Can you confirm that it is not applicable to the validation of analytical methods for residues detection. Will this guideline also be applicable to veterinary medicines? Comment: Some sections should be harmonized with FDA Guidance for Industry Bioanalytical Method Validation, May 2001 and Bioequivalence Guideline from EMEA, January 2009.	Validation of analytical methods for residues detection is outside the scope of this guideline. The EMA considers harmonisation an important topic and has therefore embarked into numerous discussions in relation to the preparation of this guideline. The outcome of these interactions on harmonisation issues have been taken into account as much as possible.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Comment: This Guideline discusses no aspects of recovery. If recovery is omitted on purpose, clarify that it is EMEAs opinion that this is not needed even though it is a requirement in the FDA guidance.	Analytical method development is outside the scope of this guideline.
45-54	23	Comments: A clearer definition of matrices in this section would be useful (serum, plasma, urine, saliva and other solid tissues…) and would allow to show the difference with the draft guideline VICH topic GL49.	The introduction section has been changed. Validation of analytical methods for residues detection is outside the scope of this guideline.
48-49	45	The list of studies used to make critical decisions is incomplete. Human pharmacodynamic studies are, for instance, not listed. Phase 3 human studies would not be included according to the definition in the text. **Proposed change (if any):** Please provide a broader list of studies (suggested wording: "The results of animal toxicokinetic and human clinical studies, including bioequivalence studies, are …"	The introduction section has been changed.
52-54	21	Lines 52-54 suggest inclusion of validation for tissue analysis within the validation guidelines. For solid tissues, it is not very clear how to measure the extraction efficiency or get accuracy for the measured concentrations in solid tissue samples. One can measure only against the tissue homogenate. The solid tissue concentration measurement should therefore not require regulated bioanalysis, but concentrations could be measured for research information purpose only. **Proposed change (if any):** We suggest to specifically exclude tissue analysis from the guideline.	The introduction section has been changed. The example of solid tissues has been deleted to prevent misunderstanding.
52 - 54	24	Comment: We would like EMA to consider the introduction of the "tiered approach" for biomarker and early metabolite quantification as well as quantification in alternative matrices (e.g., tissue) where guidelines for a full validation can not be applied for scientific reasons or limitations.	Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be subject to case by case evaluation.
52-54	45	Lines 52-54 suggest inclusion of validation for tissue analysis within the validation guidelines. For solid	The introduction section has been changed. The example of

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		tissues, it is not very clear how to measure the extraction efficiency or get accuracy for the measured concentrations in solid tissue samples. One can measure only against the tissue homogenate. The solid tissue concentration measurement should therefore not require regulated bioanalysis, but concentrations could be measured for research information purpose only. **Proposed change (if any):** we suggest to specifically exclude tissue analysis from the guideline	solid tissues has been deleted to prevent misunderstanding.
53	47	Comment: Regarding the reference to solid tissues. The validation of solid tissues is not possible for many reasons including, but not limited to, issues with internal standards, controlling and validating tissue selection, handling and storage, as well as sample preparation. **Proposed:** Recommend that the guidance be clarified to the extent that the validation of a bioanalytical method that is used for the analysis of an analyte(s) within a solid matrix is feasible only if the matrix being analyzed is a homogenate of the solid tissue. Otherwise, methods used for solid tissue analysis should be qualified rather than validated.	The introduction section has been changed. The example of solid tissues has been deleted now to prevent misunderstanding.
54	11	Comment: use a more common word to replace "prospectively"	For clarity, the text has been changed into: "Acceptance criteria wider than those defined in this guideline may be used in special situations. This should be prospectively defined based on the intended use of the method.""
55	36	**SCOPE** It is requested that the terminology and definitions used for the sates of validation are aligned to the FDA guidance terminology. Thus complete becomes "full". Partial validation is reserved when modifications are made to previously validated methods. However, the	The document has been checked for consistency and wording has been changed, if applicable. It is considered that the term 'partial validation' is sufficiently clear.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		term "Partial Validation" may be deemed to be vague and ambiguous. Thus an alternative is also suggested, the term "Supplemental Validation" be introduced in cases where work is conducted to extend a previous validation. Cross validation is used to describe circumstances where more than one valid bioanalytical methods are used to generate data within the same study. Radiolabel analysis used for the partial quantitative or qualitative approaches to determine relative distributions of metabolites and parent moieties should be out of scope. The reference removed from this section to avoid confusion. If the agency believes these methods should be validated then appropriate wording should be included.	Regarding cross validation, see section 4.3 Cross validation. Radio-labelled analysis is deleted for clarity, as it is not within the scope of this guideline.
55-62	2	Comment: • The term "Requirements" is in contradiction with "recommendations" in the EXECUTIVE SUMMARY. A guideline should recommend rather than require. • Radio-labeled analysis using 14C should not be mentioned in the guideline as those analyses are never run under GLP. **Proposed change (if any):** • Add within scope, for which bioanalytical methods this guideline is valid, e.g. GC, LC, LC-MS, GC-MS, ligand binding assays, microbiological assays.	The term 'requirements' has been changed into 'recommendations'. Radio-labelled analysis is deleted for clarity, as it is not within the scope of this guideline. The proposed text is not agreed, as new methods may be developed.
55-62	7	Comment: "full validation" instead of "complete validation" should be used. Radio-labelled analysis seems to be rather out of scope of this guidance, and therefore the last sentence should	Complete validation has been changed into full validation for consistency. Radio-labelled analysis is deleted for clarity, as it is not within

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		be deleted. **Proposed change (if any):** we suggest deleting line 61-62	the scope of this guideline.
55 – 62	13	Comment: A clear definition of the assays covered by the guideline is missing. According to the team's evaluation the following types of assays are not covered by the guideline: QC assays (because covered by ICH), assays for non ex-vivo samples (in vitro product characterization assays), semi-quantitative assays e.g. ELISA and other assays including TCID50 and microneutralization assays (would need a separate definition of acceptance criteria because of the results obtained e.g. in titers), measurement of metabolites of macromolecules except in special cases (e.g. prothrombin f1, f2). In contrast, the measurement of the drug in PK and TK studies, either as protein or in form of an activity measurement, definitely fall within the scope. This should be stated clearly.	The section on the scope of the guideline has been changed to better reflect the scope of this guideline.
55-62	21	The terminology "full validation" should be used instead of "complete validation". **Proposed change (if any):** Replace "complete" by "full" within the text of the guideline.	Complete validation has been changed into full validation for consistency.
55-62	21	Radio-labelled analyses used in PK, metabolism and mass balance studies are accepted methods that do not require validation. Therefore, radio-labelled analysis methods are out of the scope of this guideline and shouldn't be mentioned here. **Proposed change (if any):** We recommend to delete the last sentence: "Some special techniques such as radio-labelled analysis methods using 14C labelled drugs, are not covered here, but even in such cases efforts should be made to apply to the principles of this guideline".	Radio-labelled analysis is deleted for clarity, as it is not within the scope of this guideline.
55-62	23	Define more clearly what the scope is. Is it only covering	The section on the scope of the guideline has been changed to

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
59-60		active ingredients and drug products in biological matrices or something else? Here is mentioned only toxico-kinetics and clinical trials, which is limiting the scope. Radio-labelled analysis methods should be completely out of scope (therefore this sentence should be deleted from this section). It is stated that "...this guideline will describe when ... cross validation may represent an appropriate alternative ... [to full]... validation ...".How should one interpret section 4.3 in regards to the statement above? **Proposed change (if any):** Clarify, either in this paragraph or in 4.3 when cross validation is an appropriate alternative to full validation.	better reflect the scope of this guideline. Radio-labelled analysis is deleted for clarity, as it is not within the scope of this guideline. Cross validation is not considered as an alternative. The text has been changed to prevent this misunderstanding.
55-62	24	Comment: We would like ask EMA for clarification if assays like, e.g., activity assays, ^{14}C assays are included in the guideline.	Radio-labelled analysis is deleted for clarity, as it is not within the scope of this guideline.
55-62	26	Scope: This section should spell out more clearly that the document relates to nonclinical regulatory toxicology studies and human pharmacokinetic studies. Also, if bioanalytical support of drug discovery is outside the scope of this guideline, it would be helpful to indicate that. **Proposed change (if any):** Clarify the scope of the document, as indicated	The section on the scope of the guideline has been changed to better reflect the scope of this guideline.
55-62	45	it should be made clear that this guidance **does not apply** to validation and application to study samples of methods used for immunogenicity testing.	It is considered that it is clear to the reader that immunogenicity testing being a very specific topic, is not within the scope of this guideline.
55-62	52	A clear definition of the assays covered by the guideline is missing. According to our evaluation the following types of assays are not covered by the guideline: QC assays (because covered by ICH), assays for non ex-	The section on the scope of the guideline has been changed to better reflect the scope of this guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		vivo samples (in vitro product characterization assays), semi-quantitative assays e.g. ELISA and other assays including TCID50 and microneutralization assays (would need a separate definition of acceptance criteria because of the results obtained e.g. in titers), measurement of metabolites of macromolecules except in special cases (e.g. prothrombin f1, f2). In contrast, the measurement of the drug in PK and TK studies, either as protein or in form of an activity measurement, definitely fall within the scope. This should be stated clearly.	
56	3	A guidance document should not provide "requirements," but should provide "recommendations." Suggestion to change wording.	The term 'requirements' has been changed into 'recommendations'.
56	9	**Proposed change (if any):** Suggest to further define the scope to exclude validation of bioanalytical methods for biomarkers which are performed on a fit-for-purpose basis.	Methods used for determining quantitative concentrations of biomarkers used in assessing pharmacodynamic endpoints are out of the scope of this guideline. This is clearly indicated now.
57-58	21	We recommend that a clearer distinction between validation and sample analysis aspects be made throughout the guideline.	Comment noted.
57-58	45	clearer distinction between validation and sample analysis aspects is needed	Comment noted.
58	28	Comment:"…eg the actual analysis of samples from toxicokinetics studies and clinical trials." : to avoid any misinterpretation, the guideline would gain to be more precise about its scope for clinical trials : bioequivalence study, phase I trial, or any kind of clinical trial which included bioanalytical work.	The section on the scope of the guideline has been changed to better reflect the scope of this guideline.
58	47	Comment: Regarding the reference to clinical trials. If the scope of the guidance is intended to include clinical trials, then the guidance should be precise in this aspect of application and provide the respective details. This is a	The scope is referring to human and preclinical studies. The section on the scope of the guideline has been changed to better reflect the scope of this guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		very important issue as far as SQA BASS is concerned as there is currently little guidance regarding the bioanalysis of human clinical samples. **Proposed:** Recommend that the guidance indicate that the recommendations described are applicable to human clinical samples as they apply to bioanalytical method validation. Additionally, recommend further distinction between GCP and GLP requirements in the guidance. The guidance should address the applicable aspects of the role of Study Director, archiving, reporting and GLP-specific requirements to distinguish from GCP requirements, particularly those that are relevant to the quality and integrity of the bioanalytical data.	- GCP/GLP aspects are covered by their respective guidelines
61-62	28	Comment: the exclusion of radioisotope assay of the scope of the guideline is welcome but the 2nd part of the sentence (line 62) should be deleted as most of the principles described in this guideline are not applicable for liquid or solid scintillation countings/radioactivity determination in biological samples. Also the exclusion should not be limited to 14C-labeled products but to any kind of radiolabeling (3H, 14C, 32P, 125I, etc) **Proposed change (if any):** "some special techniques, such as radio-labeled analysis methods using radio-labeled drugs (^{3}H, ^{14}C, ^{32}P, ^{125}I, etc), are not covered here."	Radio-labelled analysis is deleted for clarity, as it is not within the scope of this guideline.
63	36	**Legal Basis** As the guidance refers to CPMP/ICH/381/95 which is relevant to the analysis of drug substances and drug products rather than analysis of substances in biological fluids the relevance of this reference is questioned. Greater clarity and a more robust definition of which	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		procedures are expected to be within the scope of the principles of GCP are sought. Is the assay validation for those methods expected to be used for analysis of samples taken in human clinical trials subject to the principles of GCP as well as the analysis of the samples or neither It is further suggested that the guideline directly refers to the EMEA bioequivalence guideline with respect to analysis of samples taken in those particular trials	
63 - 68	24	Comment: We propose to run validations according to SOPs in GLP compliant facilities. Please harmonize with FDA approach here. We would like to ask the EMA to consider to extent the guideline to veterinary BA, BE and PK study analysis and method validation as well.	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
63-78	2	Comment: • It is strongly recommended that validations should not be performed under GLP. However, it should be performed following the principles of GLP (status quo). • Validation of bioanalytical methods and clinical sample analysis are out of scope of OECD GLP. However, it is understood that these activities are to be performed and documented in a controlled and scientifically sound way according to principles similar to GLP. **Proposed change (if any):** • Reference to GCP and guidance on validation of analytical procedures is confusing as to their relevance to bioanalytical work, and should be deleted.	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
63-78	22	Comment: Point 3 – Legal Basis There is a great variability in the EU as to the	The section on legal basis has been rewritten for clearness.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		implementation of GLP/GCP and on how to ensure compliance in bioanalytical studies in humans. Although bioanalytical studies in humans fall outside the scope of the GLPs, EU medicines agencies and sponsors do refer to GLP certification as an essential requirement. In addition, at the beginning of the GLP implementation, the scope only covered pre-clinical studies, but nowadays most of bioanalytical work related to clinical trials is performed according to GLP. Particularly, bioequivalence studies are required to be performed according to the principles of GLP worldwide. The EGA is supportive of initiatives aiming at limiting the variability of interpretation of guideline requirements. We believe this guideline could represent an opportunity to define and harmonize the principles of GLP applicable to all bioanalytical activities. To do so, we would suggest that the text clarifies the requirements to establish what medicines agencies will accept as corroboration that studies are indeed performed according to the principles of GLP.	Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
63-78	45	It is complicated for those not familiar with the cited directives. It will be helpful to add URL addresses for the cited references.	- Adding URL addresses may lead to confusion in time, as URL addresses can change after a time.
64	21	Reference is made to a point "(4)" of the "introduction and general principles" of the Annex I to Directive 2001/83 as amended. We understand that this reference is there to remind applicants that the guideline would need to be read in conjunction with other EMA, CHMP and Community guideline published by the Commission. We would suggest that stating this verbatim in the guideline would make it easier to understand for readers.	- The guideline should be read in conjunction with other EMA, CHMP and Community guideline published by the Commission.
64	23	A reference (4) is referred to after "general principles" in this line however no references are cited in the document. If this guideline is supposed to also cover validations	- 4 stands of paragraph 4 of the introduction. The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		and bio-analyses in the veterinary medicine domain a cross reference to the EMEA VICH guidelines would be appropriate. **Proposed change (if any):** Add references.	trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
69	8	Method validation has always fallen outside the scope of the GLPs. Many parts of the GLPs will not apply to method validation (e.g. all references to animal care facilities, test and control articles, etc.) **Proposed change (if any):** Remove this requirement.	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
69- 70	23	The meaning of "in accordance with" the principles of GLP is not clear; The sentence applies to both validation of methods and analysis of study samples. To be consistent with current industry and US FDA approaches, it is preferred that method validation studies are not required to be compliant with GLP. The US FDA has stated that the GLPs do not apply to validation trials to confirm the analytical methods. It is considered that the section on GLP (Section 3 Legal basis) raises more questions than it answers and appears self- contradictory. The guideline indicates that validation of methods should be conducted in accordance with the principles of GLP. There are two issues arising from this; firstly that there has been no specific requirement up to now for method validation studies to be conducted in compliance with GLP even for pre-clinical work and secondly as the following sentence states GLP would not be appropriate for clinical studies. In the UK any claim for GLP compliance for clinical samples would not be supported and could therefore considered to be false. The guideline indicates that the analysis of study samples should be performed in accordance with GLP principles. For pre-clinical studies it would, however, depend upon the intended purpose of the study as to whether a claim of compliance is to be made.	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Section 4: In this section the validation is described. The requirement for a validation report is missing. As written this report is mentioned in section 7 and it is confusing as previous sections describes the analytical runs. Consequently there can be no reference to compliance with GLP in this report. **Proposed change:** To modify sentence to clearly define that method validations should use as a basis the applicable principles of GLP, but compliance to GLP is not a requirement.	
69-70	41	Comment: GLP to be claimed for validation studies. The general direction of the industry is currently moving in the opposite direction (no claim for GLP, except for bioanalysis within multi-site, preclinical studies **Proposed change (if any):** validation should be done in a GLP compliant facility, following the principles of GLP, but not claiming GLP.	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
69-70	45	The choice to perform method validation under GLP or not should be left to the individual company. CROs tend to perform such activity under GLP, but most pharmaceutical companies do not do it that way internally. In the case of animal studies, analysis is performed under GLP only for GLP studies (e.g., DRF, MTD studies are non-GLP toxicity studies for the in-vivo part, and are analyzed also non-GLP) **Proposed change (if any):** Modify the sentence as follows: "The validation of bioanalytical methods and the analysis of study samples may be performed in accordance with the principles of Good Laboratory Practice (GLP)."	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
69-70	47	Comment: The EMEA draft guidance indicates that validations should be conducted under the principles of GLP, yet	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		also indicates that sites conducting human studies are not required to be monitored as part of a national GLP compliance program. Based upon the current guidance wording, the possibility exists that method validations will not be conducted the same way depending on whether or not the facility is subject to GLP monitoring. Without monitoring, how will it be verified that the method validations conducted at facilities that only conduct human analyses were conducted according to GLP principles? It is critical to note that facilities that conduct analyses solely on human clinical samples may not have the necessary infrastructure (e.g., QA, training in and understanding of the GLPs, etc.) to ensure these validation studies are conducted according to GLP principles. **Proposed:** Similar to the comment made for Line 58, the bioanalysis of human clinical samples requires pertinent clarification in the guidance. Given the lack of guidance associated with human bioanalytical sample analysis and the impetus for the harmonization of bioanalytical method validation practices, SQA BASS recommends that the EMEA provide clarity in the guidance, rather than waiting for other agencies (e.g., FDA) to revise their current guidance(s).	trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
69-72	21	Validation of bioanalytical methods and clinical sample analysis are out of scope of OECD GLP. However, it is understood that these activities are to be performed and documented in a controlled and scientifically sound way according to principles similar to GLP. **Proposed change (if any):** The validation of bioanalytical methods and the analysis of study samples should be performed ~~in accordance with the principles of Good Laboratory Practice (GLP)~~ *according to applicable*	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		guidelines and compliant to Standard Operating Procedures. ~~However, as human bioanalytical studies fall outside of the scope of GLP, as defined in Directive 2004/10/EC, the sites conducting the human studies are not required to be monitored as part of a national GLP compliance programme.~~	
69 - 74	3	Assay validation and the bioanalysis of clinical samples should not be performed in accordance with GLP or GCP.	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
69-74	5	Comment: The legal situation is defined: human bioanalytical studies fall outside of the scope of GLP (line 71). In contrast, the EMA guideline requires the performance of human study "in accordance" (line 70) with GLP. As a GLP certified laboratory we see a legal discrepancy between "accordance" and "compliance". It should be made clear whether or not GLP should be applied completely. This is for us as a GLP certified laboratory of importance for our inspections/discussions with the GLP authority.	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
69-74	13	Comment: Performing a bioanalytical method validation under GLP is not conform with the OECD guidelines, which require the validation of computerized systems only	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
69-74	14	Comment: Method development and validation, as well as the bioanalytical part of clinical studies fall outside the scope of GLP (OECD 1). To ensure quality, the bioanalytical laboratories are required to have written SOPs describing the validation procedures as well as study sample analysis. GLP applies only to the bioanalytical part of safety (toxicological) non-clinical studies. Therefore, not to be in contradiction to overall OECD	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		GLP rules, the following change of the guideline text is proposed. **Proposed change:** "*The validation of bioanalytical methods and the analysis of study samples should be performed in accordance with sound principles of quality assurance and according to written SOPs. Beyond that, the bioanalytical part of preclinical safety studies (toxicokinetics) and of clinical bioequivalence trials should be performed in accordance with the principles of Good Laboratory Practice (GLP) as required by Directive 2004/10/EC and the Guideline on the investigation of bioequivalence (CPMP/EWP/QWP/1401/98).*"	
69-74	15	Method development and validation of bioanalytical methods as well as the bioanalysis for clinical studies should be conducted in a regulated environment and not necessarily in a GLP compliant environment. **Proposed change:** The validation of bioanalytical methods and the analysis of study samples should be performed in accordance with sound scientific principles and well established procedures (e.g. SOPs) to guarantee quality of the results. The principles of Good Laboratory Practice as required by Directive 2004/10/EC should be applied to the bioanalytical part of preclinical safety studies (toxicokinetics). Bioanalysis in the framework of clinical bioequivalence trials should be performed in accordance with the principles of Good Laboratory Practice (GLP) as required by the Guideline on the investigation of bioequivalence (CPMP/EWP/QWP/1401/98).	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
69-74	28	Comment: For preclinical assays (TK), it is understood that now compliance to GLP will be mandatory for both	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		the validation and the assays. This inclusion of the validation in the GLP scope is welcomed as the current situation was very confusing in our opinion (validation not mandatory to be GLP-compliant and assays GLP-compliants). However, for clinical trials, the GLP & GCP scope proposed in the guideline is very confusing. It is understood that validation & assays in clinical trials will need to be GLP-compliant but that this analytical work can be done in facilities not followed for GLP compliance by Authorities. **Proposed change (if any):** either all the facilities conducting bioanalytical work (for non-clinical AND clinical trials) should be GLP compliant AND part of the GLP compliance monitoring programme, or validation/assays of clinical samples should not be mandatorily done in compliance to GLP.	Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
69-74	44	Comment: The GLP Principles should be followed. In that case, the test sites should be subjected to inspections. Otherwise claims of GLP adherence are not verified. This leaves the way open to false claims and unverifiable integrity of data. Directive 2004/10/EC, Article 1, Par. 1: "Member States shall take all measures necessary to ensure that laboratories carrying out tests on chemical products, in accordance with Directive 67/548/EC, comply with the Principles of GLP …..". This applies also where other Community provisions provide for the application of GLP (Article 1, second paragraph). Directive 2004/10/EC does not define the non-clinical part of bioanalytical studies and bioanalytical validation studies as being outside of the scope of GLP. Directive 2004/9/EC applies to non-clinical safety testing on all types of chemicals (Art. 1); Member States shall verify GLP compliance of any testing laboratory claiming	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		to use GLP (Art. 2). Monitoring programmes of some Member States may not allow for inspections of GLP compliance of bioanalytical laboratory work that is part of human studies. **Proposed change:** The validation of bioanalytical methods and the analysis of study samples should be performed in accordance with the principles of Good Laboratory Practice (GLP). Sites conducting the analysis of samples and/or the validation of bioanalytical methods should be allowed to be monitored as part of a national GLP compliance programme. In addition, for clinical trials in humans the principles of Good Clinical Practice (GCP) should be followed.	
69-74	52	Performing a bioanalytical method validation under GLP is not conform with the OECD guidelines, which require the validation of computerized systems only	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
74	33	It could be advisable to add two lines (or a short new paragraph) to remind all interested parties that: " The guideline ICHQ9 - Quality Risk Assessment - strongly recommends that the validation of analytical methods be performed using previously validated (critical) instruments and executed by properly trained technicians". We suggest to add these lines in Paragraph 3 "Legal Basis " at line 74.	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any misunderstanding regarding this issue.
76	47	Comment: Regarding the reference to the EMEA guideline CPMP/ICH/135/95.	The section on legal basis has been rewritten for clearness. Furthermore, reference is made to the "Reflection Paper for Laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) to prevent any

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed: Recommend removing as the referenced document contains no relevant information to the bioanalytical analyses of samples from clinical trials.	misunderstanding regarding this issue.
79 - 82	24	Comment: Full validation for each species should not be required for chromatography based assays. Please consider the use of partial validation in line with FDA guidance. Please add serum to the list of matrices. For alternative matrices we would like EMA to consider our comments above for the tiered approach for alternative matrices.	The proposed reference to define which issues should be addressed in the partial validation cannot be confirmed. Serum has been added to the list of matrices. Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be subject to case by case evaluation.
79-98 line 86	36	"A full validation is required in each species concerned" This is contrary to CDER BMV guidelines where a partial validation is acceptable in other species following a full validation in the first toxicology species. **Proposed change (if any):** An abbreviated (partial) validation should be recommended with a justification for the use of a partial validation	- This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc..
79 - 316	3	The general recommendation from the responding companies is to organize the guideline into separate sections to provide clarity to the user -- one section dedicated to LC-MS method validation and sample analysis and a separate section dedicated to LBA method validation and sample analysis. As written, the LBA reader particularly must flip back and forth to piece together all the required elements of a LBA validation. LBAs employ unique critical reagents. The guidance should include a discussion on the relevant considerations needed, particularly in regard to the impact of changing critical reagents during validation or sample analysis.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
79 - 316	3	The acceptance criteria stated in **Section 4. Method Validation** are specific to chromatographic assays and	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		not relevant or appropriate for LBAs. Thus, the guidelines for each parameter (specifically Selectivity, Dilution integrity, Matrix effect, Stability, Partial validation, Cross validation)and the relevant acceptance criteria should be defined for LBAs.	recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
79 - 316	3	A section should be added to **4. Method Validation** to discuss Ruggedness and Robustness. **Ruggedness and Robustness** – both parameters are indicators of how reproducible the method remains under real life changes in the laboratory environment, either pre-meditated or by chance. Robustness refers to the ability of the method to maintain validity when small changes to the specific method are applied, e.g., incubation times, light. Ruggedness refers to the methods consistency after changes, external to the specific method, are made like change in analyst, or equipment.	Ruggedness and robustness are considered issues for method development. Furthermore, it is difficult to address the criteria to be evaluated for these issues.
79-98	2	Comment: • The term "complete" should be avoided when associated with validation as a "full validation" could be suggested. A clear definition of full, partial and cross validation should be provided. • We request clarification regarding the need to demonstrate internal standard stability in the matrix. Internal standards are not stored with samples and, therefore, stability determination is not required. IS should in general provide a reasonable uniform response throughout the run, but that too is under debate (see 'Validation topics with no consensus' in Crystal city 3 whitepaper, AAPS Jr. 2007: 9(1), article 4). **Proposed change (if any):** • Please remove or rephrase "same anticoagulants as for the study samples". Partial validation is the recommended approach when the anticoagulant in	Complete validation has been changed into full validation for consistency. Stability of the internal standard in matrix has not to be evaluated. The text has been changed accordingly. As indicated, validation should be carried out using the same anticoagulant as used in the study samples.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		the study and validation samples differs.	
80	21	We would suggest to indicate that section 4.1 is dedicated to physico-chemical assays and since there is a specific section 4.4 for ligand-binding assays.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
80	23	Indicate that this section 4.1 is decdicated to physico-chemical assays and that there is a specific section or guideline for ligand-binding assays.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
80	45	**Proposed change (if any):** Replace "complete validation" by "full validation"	Complete validation has been changed into full validation for consistency.
80-98	22	Comment: Point 4.1 - Complete validation of an analytical method The EGA supports the principle that a complete method validation should be performed for any analytical method whether new or based upon literature when it is firstly implemented in a facility. However, it should be acceptable in certain defined and justified situations, to use cross-validation or re-validation (with a reduced validation protocol, i.e. partial validation). An example of such situation would be a change of anticoagulant for an anticoagulant of the same family (eg, EDTA K2 or K3 with only a difference in the number of counter-ions) where we believe a partial validation should be considered sufficient and that new stability studies should not be required. The EGA would welcome clarification on this point.	For anticoagulants only differing in counter ion, a partial validation may be applicable. However, this issue has been discussed before at several meetings, but no consensus could be obtained. Furthermore, it is known for EDTA K2 and EDTA K3, that a difference in pH can occur of the matrix which may effect the analytical methods (stability for instance; see Asanuma et al., Lab. Hematology 6:67-72 (2000) and NCLLS, Additives to blood collection devices: EDTA; H35-T (ISBN 1-56238-162-8)). This issue is subject to a case by case situation. .
80-299 section 4.1	47	Comment: The method parameters described throughout Section 4.1 (Complete validation of an analytical method) fully	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		addresses parameters for LC/MS and provides only limited guidance for immunoassays/ligand-binding assays. **Proposed:** There is an immediate need for a guidance document that would extend beyond a stated understanding that the basic principles of validating bioanalytical methods for small molecule also apply to ligand-binding assays, except for (1) variation in the types of validation method characteristics that require investigation and documentation, and (2) allowance for looser acceptance criteria because of the unique heterogeneous nature of macromolecules and other aspects of the immunoassay (supported with scientific rationale). SQA BASS feels that the EMEA guidance offers the ideal opportunity to provide guidance relevant to immunoassay/ligand-binding assays since having such a guidance would facilitate consistent regulatory approaches to ensuring adequate practices, such as scientific peer review and QAU/agency auditing, while allowing for the creative scientific decisions necessary to facilitate effective drug development and appropriate drug submissions. It is the SQA BASS recommendation that a similar approach and language to immunoassays (at least ligand-binding assays) is needed throughout Section 4.1, including the necessary additional required experiments associated with ligand-binding assays, since the approach to validating both techniques is similar. If there are EMEA concerns regarding fluidity of text in the guidance document, recommend that an additional section(s) be included to address ligand-binding assays (and ideally other immunoassays and biomarker assays) and be structured similar to Section 4.1.	separate section has been introduced on ligand binding assays.
81	8	Comment: No requirement for complete validation when another analyte is added to the method (refer to FDA guidance document). Should this be added? **Proposed change (if any):** Add this requirement for	Adding another analyte to the matrix can be regarded as a difference in matrix. This would trigger evaluation of for instance matrix effect, selectivity, precision/accuracy, and stability. However, no specific guidance is added regarding this issue and should be considered as a case by case

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		harmonization.	evaluation.
83-86	7	Comment: In case the same analytical method is used for various toxicology species, it should be possible to limit a full validation to the first toxicology species and allow a partial validation for subsequent species. Of course, the partial validation would cover areas depending on the species such as matrix stability but would exclude species independent experiments such as stock solution stability. The expression "same anticoagulant" should allow some flexibility, i.e. the use of K2 and K3 EDTA without the need for a cross-validation.	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc.. For anticoagulants only differing in counter ion, a partial validation may be applicable. However, this issue has been discussed before at several meetings, but no consensus could be obtained. Furthermore, it is known for EDTA K2 and EDTA K3, that a difference in pH can occur of the matrix which may effect the analytical methods (stability for instance; see Asanuma et al., Lab. Hematology 6:67-72 (2000) and NCLLS, Additives to blood collection devices: EDTA; H35-T (ISBN 1-56238-162-8)). This issue is subject to a case by case situation.
83-86	21	In case the same analytical method is used for various toxicology species, it should be possible to limit a full validation to the first species and allow a partial validation for subsequent species. Of course, the partial validation would cover areas depending on the species such as matrix stability but would exclude species independent experiments such a stock solution stability.	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc..
84-85	45	**Proposed change (if any):** Add serum as matrix.	Serum has been added to the list of matrices.
84-85	47	Comment: Regarding the reference to validation of methods for urine and tissue matrices, validating these types of matrices is not realistic unless the collection and stability prior to freezing can be regulated at the site of	The section on introduction and the scope of the guideline has been changed to better reflect the scope of this guideline. The example of solid tissues has been deleted to prevent

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		collection. **Proposed:** Recommend that the collection, storage stability, stability during collection, and preparation, buffering, and/or volumes be regulated to the point that the applicable bioanalytical method can be appropriately validated. Recommend clarification of these considerations in the guidance. Similar to the comment for Line 53, validation of certain tissue methods may not be feasible.	misunderstanding. The proposal has not been included as it is considered that the issue is resolved by changes in the introduction section and scope section.
85	8	Comment: When referring to the "same anticoagulant" in this line and throughout document, it is unclear if the same counter-ion is also required. Industry is currently being questioned on this topic, and the guidances are unclear on the requirement. **Proposed change (if any):** Clarify if the same anticoagulant counter-ion is required between method validation and sample analysis.	For anticoagulants only differing in counter ion, a partial validation may be applicable. However, this issue has been discussed before at several meetings, but no consensus could be obtained. Furthermore, it is known for EDTA K2 and EDTA K3, that a difference in pH can occur of the matrix which may effect the analytical methods (stability for instance; see Asanuma et al., Lab. Hematology 6:67-72 (2000) and NCLLS, Additives to blood collection devices: EDTA; H35-T (ISBN 1-56238-162-8)). This issue is subject to a case by case situation.
85	9	**Proposed change (if any):** Suggest to add "serum" to this list as an example, as it is the most frequently used matrix for ligand binding analysis.	Serum has been added to the list of matrices.
85	23	"..validation should be performed using the same anticoagulant as for the study samples." why use the word "should "if the same anticoagulant has to be used? Does this mean that a new complete validation must be performed if the anticoagulant is changed, or is a partial validation sufficient? The same question arises for the change in matrix. **Proposed change:** Validation has to be performed using the same anticoagulant as for the study samples."	For anticoagulants only differing in counter ion, a partial validation may be applicable. However, this issue has been discussed before at several meetings, but no consensus could be obtained. Furthermore, it is known for EDTA K2 and EDTA K3, that a difference in pH can occur of the matrix which may effect the analytical methods (stability for instance; see Asanuma et al., Lab. Hematology 6:67-72 (2000) and NCLLS, Additives to blood collection devices: EDTA; H35-T (ISBN 1-56238-162-

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Clarify the EMEA standpoint regarding K2-EDTA/K3-EDTA should be considered same or different anticoagulant. **Proposed change (if any**): Include an example in the sentence.	8)). This issue is subject to a case by case situation.
85	26	Anticoagulant is most likely only used in plasma collection **Proposed change (if any):** Insert ", if an anticoagulant is used," before ", validation should be …"	This is agreed. The text has been changed accordingly.
85	45	We do not support validation of methods for tissue analysis: too many uncertainties! In addition, we do not support validation of methods for changes to different strains or genetic pools	The section on introduction and the scope of the guideline has been changed to better reflect the scope of this guideline. The example of solid tissues has been deleted to prevent misunderstanding. Regarding different strains and genetic pools, the guideline does not request validation methods in these cases.
85-86	21	**Proposed change (if any):** We would suggest that if changes between anticoagulant occurs this must only be partially validated	For anticoagulants only differing in counter ion, a partial validation may be applicable. However, this issue has been discussed before at several meetings, but no consensus could be obtained. Furthermore, it is known for EDTA K2 and EDTA K3, that a difference in pH can occur of the matrix which may effect the analytical methods (stability for instance; see Asanuma et al., Lab. Hematology 6:67-72 (2000) and NCLLS, Additives to blood collection devices: EDTA; H35-T (ISBN 1-56238-162-8)). This issue is subject to a case by case situation.
85-86	37	Comment: Risk-based approach and practical aspects should be taken into account: the biological material is complex in nature and consists of numerous unidentified substances of different identity and quantity in different lots of material. It seems unreasonable to concentrate	For anticoagulants only differing in counter ion, a partial validation may be applicable. However, this issue has been discussed before at several meetings, but no consensus could be obtained.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		on anticoagulant, which is usually of different chemical structure than analyte and may be lost during sample preparation, chromatographic separation as well as may be undetectable in analysis conditions. There is no or very little evidence in literature that presence/absence of anticoagulant is critical for measurement results reliability. Generally, as described in lines 468-471, the validation samples do not mimic the actual study samples. Moreover, as described in lines 87-89, EMA allows synthetically prepared cerebrospinal fluid. How dose it correspond with strict regulation on anticoagulant? Is there reliable scientific evidence that anticoagulant influence method selectivity, precision/accuracy or analyte stability? In the case it is not acceptable to delete the anticoagulant requirement, it is suggested that partial validation of clearly stated range of experiments should prove method reliability when anticoagulant is changed. **Proposed change (if any):** ~~Moreover, validation should be performed using the same anticoagulant as for the study samples.~~	Furthermore, it is known for EDTA K2 and EDTA K3, that a difference in pH can occur of the matrix which may effect the analytical methods (stability for instance; see Asanuma et al., Lab. Hematology 6:67-72 (2000) and NCLLS, Additives to blood collection devices: EDTA; H35-T (ISBN 1-56238-162-8)). This issue is subject to a case by case situation.
85-86	42	Comment: *Moreover ,validation should be preferably performed using the same anticoagulant* (Since it is very difficult to get the same type of anticoagulant)	As indicated, validation should be carried out using the same anticoagulant as used in the study samples. It is not agreed that it is very difficult to obtain the same type of anticoagulant, however it may be more costly in certain cases. For anticoagulants only differing in counter ion, a partial validation may be applicable. However, this issue has been discussed before at several meetings, but no consensus could be obtained. Furthermore, it is known for EDTA K2 and EDTA K3, that a difference in pH can occur of the matrix which may effect the analytical methods (stability for instance; see Asanuma et al., Lab. Hematology 6:67-72 (2000) and NCLLS, Additives to blood collection devices: EDTA; H35-T (ISBN 1-56238-162-

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			8)). This issue is subject to a case by case situation.
85-86	45	**Proposed change (if any):** To specify anticoagulant might be too specific, if changes between anticoagulant occurs it must be partially validated	As indicated, validation should be carried out using the same anticoagulant as used in the study samples. For anticoagulants only differing in counter ion, a partial validation may be applicable. However, this issue has been discussed before at several meetings, but no consensus could be obtained. Furthermore, it is known for EDTA K2 and EDTA K3, that a difference in pH can occur of the matrix which may effect the analytical methods (stability for instance; see Asanuma et al., Lab. Hematology 6:67-72 (2000) and NCLLS, Additives to blood collection devices: EDTA; H35-T (ISBN 1-56238-162-8)). This issue is subject to a case by case situation.
85-87	49	There is no mention of the counter-ion. **Proposed change (if any):** Same anticoagulant and counter-ion should be used for validation and study samples	For anticoagulants only differing in counter ion, a partial validation may be applicable. However, this issue has been discussed before at several meetings, but no consensus could be obtained. Furthermore, it is known for EDTA K2 and EDTA K3, that a difference in pH can occur of the matrix which may effect the analytical methods (stability for instance; see Asanuma et al., Lab. Hematology 6:67-72 (2000) and NCLLS, Additives to blood collection devices: EDTA; H35-T (ISBN 1-56238-162-8)). This issue is subject to a case by case situation.
86	14	Comment: In order to save animals life (e.g. rodents), to reduce blood sampling in animals and to consider animal welfare laws, the amount of plasma or other matrices needed for validation purposes should be as small as possible. Therefore, a complete validation should be	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc..

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		limited to one animal species. For further animal species partial validations should be considered as an alternative approach if reasonable quality can be derived from the complete validation. **Proposed change:** The wording "*A full validation should be performed for each species concerned*" should be changed to "*To reduce blood sampling in animals, a complete validation should be limited to one animal species. For further animal species partial validations should be considered as an alternative approach if reasonable quality can be derived from the complete validation. However, for human bioanalytical methods always a complete validation is required.*"	It is recognised that this aspect addresses the ethical aspect against the certainty/reliability aspects.
86	23	"A full validation is required for a change of species". What is the scientific rationale justifying to validate all the parameters required for a full validation. What would be the utility to perform stability of the stock and working solutions? We believe that a full validation is not necessary for each individual species: e.g. stability of extracted samples, influence of haemolysis, dilution effect, etc.: as far as it has been established in one species, there is very reduced chance that it would differ between species. Even the matrix variability, which is critical for human species, has less chance to be impacted in preclinical species as the animals come from same strains and have similar food. **Proposed change:** The compromise would be to perform a partial validation focused on the differences linked to the species as it is described in the report of the conference "Bio-analytical Method Validation – a revisit with a decade of progress" Workshop held in Arlington, VA, 2000. The partial validation would include selectivity, long-term stability, within-run accuracy and precision.	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc.. It is recognised that this aspect addresses the ethical aspect against the certainty/reliability aspects.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
86	28	Comment: "a full validation should be performed for each species concerned" : for some techniques, a partial validation may be considered when the different species/same matrice are studied (see comment below)	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc.
86	34	Full validation should be performed for the main tox species and for humans, other studies conducted in different species or different matrices could be done using a partial validation	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc.. It is recognised that this aspect addresses the ethical aspect against the certainty/reliability aspects.
86	37	Comment: it is suggested that partial validation of clearly stated range of experiments should prove method reliability in case of switch between plasma and serum of one specimen. **Proposed change (if any):** In case of switch between plasma and serum of one specimen, only partial validation is required.	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc.. It is recognised that this aspect addresses the ethical aspect against the certainty/reliability aspects.
86	41	Comment: Full validation required for each species is contradicting the FDA Guidance, in which only a full validation is required for the first species and for human. If the assay is similar, only a partial validation can be applied. **Proposed change (if any):** Make in line with the FDA Guidance.	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc.. It is recognised that this aspect addresses the ethical aspect against the certainty/reliability aspects.
86	45	Partial validation is listed in FDA guidance as acceptable for change in species for the same matrix (e.g. for change from rat serum into monkey serum). See also Lines 300-307 (partial validation). We suggest that relevant experiments be performed for new species, a full validation may not be necessary	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc.. It is recognised that this aspect addresses the ethical aspect

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			against the certainty/reliability aspects.
87-89	21	For ELISA methods the statement is true. However, for chromatography based methods, combination of full validation for one species and partial validation to extend to another matrix or species results in a fully validated method for the old and the new matrix, species, anticoagulant etc.	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc.. It is recognised that this aspect addresses the ethical aspect against the certainty/reliability aspects.
87-89	45	For ELISA it is true. For chromatography based methods, combination of full validation for one species and partial validation to extend to another matrix or species results in a fully validated method for the old and the new matrix, species, anticoagulant etc.	This is partially agreed. For example stock solutions stability has not to be evaluated again. However, matrix evaluations are necessary. It cannot be foreseen on forehand whether difference in matrix with different endogenous compounds does not affect LLOQ, stability, precision etc.. It is recognised that this aspect addresses the ethical aspect against the certainty/reliability aspects.
88	23	When will it be possible to use synthetic matrixes (such as synthetic urine)? What about matrix to use for endogenous compound?	There are always exceptions which cannot be covered by this guideline. This is a case by case situation.
88-89	47	Comment: Regarding the reference to alternative matrices (surrogate) and consideration of original matrix for QC samples when surrogate matrix is used for the calibration curve. **Proposed:** Recommend that the use of a surrogate matrix is applicable to calibration curve standards, but not QC samples. Further recommend that it be required that matrix QC samples be used for the method validation and subsequent study sample analyses, as opposed to using QC samples prepared with a surrogate or artificial matrix.	There are always exceptions which cannot be covered by this guideline. This is a case by case situation. In case of limited availability it is obvious that an alternative has to be used. In addition, this may also applicable to the use of study samples as QC samples, however, there is still the problem that the actual concentration in the study sample is not known in case of endogenous compounds.
88-89 and	41	Comment: The description in these lines is contradicting	The use of an alternative matrix is an exception and not a

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
156-157		– use of alternative matrix (line 88-89) vs. same matrix (line 156-157). **Proposed change (if any):** make in line with each other. Please note that the use of the same matrix is not possible for endogenous analytes.	general rule.
90-94	7	Comment: The stability of the internal standard in the matrix is not relevant because the internal standard is not part of the study samples. **Proposed change (if any):** Delete ... "and any internal standard in the biological matrix" to read as follows: "The main characteristics of a bioanalytical method that are essential to ensure the acceptability of the performance and the reliability of analytical results are: selectivity, lower limit of quantitation, the response function (calibration curve performance), accuracy, precision, matrix effects, stability of the analyte(s) ~~and any internal standard in the biological matrix~~ and the stock and working solutions under the entire period of storage and processing conditions."	Stability of the internal standard in matrix has not to be evaluated. The text has been changed accordingly.
90-94	19	Recovery is to be evaluated according to the FDA guideline (2001). **Proposed change (if any):** We suggest to include this test to be in line with FDA requirements.	Recovery is an issue investigated during the analytical method development and as such considered not necessary to be included in this guideline.
90-94	21	We would argue that the stability of the internal standard in the matrix is not relevant since the IS is not part of the study samples and that stability assessment of internal standard would only be of limited value. This can be addressed in method development as suggested in the FDA guidance, or even not at all. No certificate of analysis for IS should be required. That is not entirely clear from the text of the draft guideline and should be clarified. **Proposed change (if any):** We would recommend to delete ... "and any internal standard in the biological	Stability of the internal standard in matrix has not to be evaluated. The text has been changed accordingly.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		matrix"...in line 93	
90-94	21	Dilution parallelism and linearity for ELISA is missing as well as interference from e.g. soluble receptors and soluble targets.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
90 - 94	24	Comment: Stability of the internal standard (IS) is not an issue in biological matrices, such stability data are generated during method development. For LBA we would like to see as well dilution integrity, parallelism, hook effect and interference from e.g., soluble receptors and targets as additional method parameters.	Stability of the internal standard in matrix has not to be evaluated. The text has been changed accordingly. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
90-94	26	The calibration range is also important to define **Proposed change (if any):** Insert calibration range as a characteristic	Calibration range has been included in the text.
90-94	45	Dilution parallelism and linearity for ELISA is missing as well as interference from e.g. soluble receptors and soluble targets.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
90-94	47	Comment: Regarding the listing of characteristics essential to method validation, the concept of recovery experiments, an industry standard practice and discussed in the Crystal City meetings as well as the FDA guidance on method validation has not been discussed anywhere in this guidance. Additionally, there does not appear to be consideration given to characteristics associated with immunoassays, including ligand-binding assays (LBAs), such as cut point, parallelism, etc.	Recovery is an issue investigated during the analytical method development and as such considered not necessary to be included in this guideline. It is recognised the fact that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed: Recommend that guidance on recovery of the analyte(s) of interest as well as internal standard(s) is discussed within this document. It is essential that all method types and associated acceptance criteria for ligand-binding assays be addressed. Additionally, those associated with other immunoassay methods should be addressed if it is the intention of the EMEA to include these in the scope of this guidance. Recommend that the context of the document treat LBA and LC/MS acceptance criteria similarly, even if there are circumstances with LBA methods that are more variable (and thus allow for greater flexibility in acceptance parameters). This is already considered a "best practice" in industry, and thus the guidance should reflect this, especially since the technologies and methodologies associated with immunoassays are advancing as rapidly as the complexity of the associated macromolecules.	
90-94	49	There is no mention of recovery. Is recovery required?	Recovery is an issue investigated during the analytical method development and as such considered not necessary to be included in this guideline.
90 - 98	24	Comment: We would like EMA to consider here as well the use of the tiered approach for additional analytes.	Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be subject to case by case evaluation.
91	23	There is no more need to evaluate the recovery?	Recovery is an issue investigated during the analytical method development and as such considered not necessary to be included in this guideline.
91-92	36	Guidance suggests that it is essential to ensure the acceptability of the performance includes response function. This may not be applicable for Ligand Binding Assays (LBA) Clarification should be sought or added to indicate that certain parameters are only applicable only for LC-MS/Ms based platforms and that other assay platforms	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		may require alternative wording and clarity	
91 - 94	20	Comment: Acceptable dilution validation data are essential to report sample data with analyte concentrations > ULOQ **Proposed change (if any):** Recommend including dilution validation to the list	The main characteristics that are considered essential are stated. Dilution is only needed in case study samples have been diluted.
90-94	22	Comment: Point 4.1 Reference standards A reference to the methodology to be used for recovery is missing. **Proposed change (if any):** We would appreciate clarification as to the recovery methodology to be used.	Recovery is an issue investigated during the analytical method development and as such considered not necessary to be included in this guideline.
91, 123, 132, 150, 151, 162, 547, 568, 569, 589 and 590	4	Comment: Line 152 uses the verb quantified. We would prefer the noun derived from this verb (quantification) rather than the American quantitation, which is used extensively in the document. **Proposed change (if any):** Replace quantitation with quantification.	'Quantitation' has been replaced by 'quantification' in this guideline.
92-93	14	Comment: Investigations on stability of internal standard (IS) are not required since the IS will be spiked freshly at a constant amount to calibration-, QC- and study samples. Especially in the case of labelled IS(s) stability investigations on IS(s) are redundant with the stability investigations of the analyte(s). **Proposed change:** Please delete that the stability of IS should be investigated.	Stability of the internal standard in matrix has not to be evaluated. The text has been changed accordingly.
92-94	45	Internal standards are not stored with samples and therefore stability determination is not required. IS	Stability of the internal standard in matrix has not to be evaluated. The text has been changed accordingly.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		should in general provide a reasonable uniform response, but that too is under debate (see 'Validation topics with no consensus' in Crystal city 3 whitepaper, AAPS Jr. 2007: 9(1), article 4).	
93	11	Comment: usually no need to evaluate stability of internal standard in matrix	Stability of the internal standard in matrix has not to be evaluated. The text has been changed accordingly.
93	36	Text appears to imply that stability of the internal standard in the biological matrix is required. This is not a common practice in validation and can be addressed in method development. Clarification around this is requested.	Stability of the internal standard in matrix has not to be evaluated. The text has been changed accordingly.
93	39	It should be emphasized that an internal standard is normally added to the samples during sample processing. Thus, it has no sense to test its stability in the matrix. **Proposed change (if any):** Delete the words "and any internal standard".	Stability of the internal standard in matrix has not to be evaluated. The text has been changed accordingly.
93-94	23	Should the stability of stock and working solutions be included in each "full validation" or can this be information be taken from other validations using these same materials which would have no effect on these parameters?	This is considered acceptable, as long as these data are obtained from similar storage conditions.
93 and 265	16	Comment: We request clarification regarding the need to demonstrate internal standard stability in the matrix. The frequent use of stable label internal standards for mass spectrometry would obviate the need for retesting the same chemical moiety a second time. **Proposed change (if any):** Please limit testing the stability of internal standards to chemical analogs	Stability of the internal standard in matrix has not to be evaluated. The text has been changed accordingly.
95-98	7	Comment: the tiered approach should be considered depending on the status of the study.	Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be subject to case by case evaluation.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
95-98	21	We would recommend to consider a tiered validation approach for urine, tissue etc. Method validation should be limited to those matrices for which either PK analysis is the primary endpoint OR for those cases where safety is concerned (e.g. toxicity target tissue types). For all other situations we would recommend to consider using a qualified assay approach.	Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be subject to case by case evaluation. The introduction section and scope has been changed. The example of solid tissues has been deleted to prevent misunderstanding.
95-98	23	The use of examples can be improved with e.g. microorganisms used as active ingredients. Do analytes also include provoked responses like antibodies or other biomarkers?	Methods used for determining quantitative concentrations of biomarkers used in assessing pharmacodynamic endpoints are out of the scope of this guideline. This is also the case for antibodies and active micro organisms.
95-98	36	**Comments.** Clarity is sought as to how this may apply to biologics. For example measurements of free and bound forms of the compound (e.g. monoclonal antibody soluble target) may also be done.	Methods used for determining quantitative concentrations of biomarkers used in assessing pharmacodynamic endpoints are out of the scope of this guideline. This is also the case for antibodies.
95-98	36	**Comments.** Clarity is sought on how assays (e.g. Multiplex LBA assays) are to be validated	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
98	8	Comment: There is no mention of treating analytes as distinct during validation (e.g. how do we treat a stability batch that meets criteria for the drug but not the metabolite?).	Analytes should be validated together, but the results evaluated separately.
99-117	2	**Reference standards** Comment: • Unlabeled analyte will always be present to some extent in a stable labeled IS and, therefore, this fact should be handled such that labeled IS can be used without introducing a bias in the assay. **Proposed change (if any):**	The proposed text has been taken into account and line 100-102 has been changed accordingly. Regarding line 116-117, the proposed text has been partially incorporated.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		• Change lines 100-102 to the following: "During method validation, a blank biological matrix will be spiked with the analyte(s) of interest using solutions of reference standard(s) to prepare calibration standards, QC and stability samples. In addition, suitable internal standard(s) is/are added during sample processing in chromatographic methods." • Change last sentence in lines 116-17 to the following: "The presence of any unlabeled analyte should be checked to show that it will not introduce a bias in the results"	
99 - 117	3	Note the different challenges of LBA reference standard which are typically non-homogeneous and in liquid form. The discussion for reference standard for large molecules should center on the fact that many protein biotherapeutics are produced in cells and thus, just like in the body, have a range of acceptable and active forms. As a direct result of how macromolecules are produced, the reference standards tend to be heterogeneous, *vis-à-vis* post-translational modification (e.g., glycosylation or phosphorylation), 3-dimensional structure and folding. In contrast, small molecule reference standards are homogeneous with a high degree of purity. Suggest discussion on change of lot (CMC specification requirement) versus change in expression system (considered a new entity possibly).	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
99-117	22	Comment: Point 4.1 Reference standards The requirements applicable to the Internal Standards (eg, characteristics, documentation) are currently unclear in the draft guideline text. The EGA believes that all reference standards of the analytes, regardless of	It is considered that for an internal standard a Certificate of Analysis is not necessary, as long as the suitability for use as internal standard has been proven.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		their origin, but also Internal Standards, should be accompanied by a Certificate of Analysis. In the case of bioequivalence studies, the commonly used isotopes are deuterated isotopes. These are well-known for their stability and the absence of isotope exchange reaction. We would therefore recommend deleting the requirement to demonstrate the absence of exchange reactions in the specific case of well-known stable isotopes.	The recommendation of deleting the requirement to demonstrate the absence of exchange reactions is disagreed, as the guideline does not indicate demonstration of the absence.
99-117 Reference Standards	36	**Comments** Clarity is requested on the terminology of "blank biological matrix" and how this applies to pooled or individual spiked samples.	Reference is made to section 4.1.1 Selectivity.
101-102	23	"An internal standard is normally used". Keep the possibility to not use an Internal Standard. Do the standards also include e.g. microorganisms or nucleic acids?	The text has been changed into 'In addition, suitable internal standard(s) (IS) can be added during sample processing in chromatographic methods.'. This keeps the possibility open not to use an internal standard. Microorganisms and nucleic acids are not within the scope of this guideline.
101-102	39	It is true that in the majority of methods an IS is used and in LC-MS methods it is necessary to use it, but in methods with optical detection and simple sample preparation external standardisation can be used sometimes with better precision or accuracy. **Proposed change (if any):** Change the sentence to: In addition, an internal standard (IS) is normally used in chromatographic methods, but external standard method can be also used.	The text has been changed into 'In addition, suitable internal standard(s) (IS) can be added during sample processing in chromatographic methods.'. This keeps the possibility open not to use an internal standard.
101-102	42	Comment: **Proposed change (if any):** For non-chromatographic methods, the addition is Internal standard is not needed	The text has been changed into 'In addition, suitable internal standard(s) (IS) can be added during sample processing in chromatographic methods.'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		(This is needed, when we analyse the iron, lead etc by non-chromatographic methods, like AAS ICP-plasma analyzer)	This keeps the possibility open not to use an internal standard. Furthermore, the text does not imply that this is the case for non-chromatographic methods.
103 – 104	13	Comment: For biological assays the purity of the reference standard is not as obligatory as for chromatographic methods, for which e.g. an HPLC grade standard is required. For biological assays the potency is the parameter to consider (see general remarks). The reference standard is critical for the assay, most of the time the reference standards are provided by the vendors, however, the "well characterized in house material" sometimes could also be used as reference standard. It is proposed that this document should include a section to specify the requirements to establish the reference standard stability (for shelf life or the expiration date). The document outlined the "sample stability" establishment, that section may be expanded to cover the reference standard expiration date establishment requirements.	Biological assays are not under the scope of this guideline.
103-104	52	For biological assays the purity of the reference standard is not as obligatory as for chromatographic methods, for which e.g. an HPLC grade standard is required. For biological assays the potency is the parameter to consider (see general remarks). The reference standard is critical for the assay, most of the time the reference standards are provided by the vendors, however, the "well characterized in house material" sometimes could also be used as reference standard. It is proposed that this document should include a section to specify the requirements to establish the reference standard stability (for shelf life or the expiration date). The document outlined the "sample stability" establishment, that section may be expanded to cover the reference standard expiration date establishment requirements.	Biological assays are not under the scope of this guideline.
103-111	21	Metabolite characterization is only possible when available amounts of reference standards allow. In other	Some compounds used as internal standard or rare metabolites are available in very small amounts and their

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		situations, e.g. exploratory studies, exemption for complete characterization should be acceptable.	Certificate of Analysis may be not available. If the full Certificate of Analysis is not available in such cases, at a minimum the documented purity information should be obtained.
103 - 111	24	Comment: Please define the term fully characterized as this give terminology relates to GMP and is not clearly understood in the context of bioanalysis. For LBAs CoA is needed for both, validation and study conduct. **Proposed change (if any):** We would propose to align with FDA. Comment: The presence of any unlabelled analyte in stable isotope-labeled IS material will not generally bias the results. It rather influences the LLOQ of the method. It should be investigated as well that the drug analytes doesn't interfere with the stable labeled internal standard and thus impact the result. **Proposed change (if any):** Therefore, delete the sentence or replace the sentence by: "If relevant amounts of unlabelled analyte are detected the potential influence has to be evaluated during method validation."	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. For chemically entities purity and identity is considered sufficiently. For LBA, i.e. macromolecules also fully characterisation of the structure to the extent possible is needed. . The proposed change is partially incorporated in the text.
106	36	**Comments** Clarity is sought as to what "fully characterized" represents. For example many biological reference standards characterization may be more limited than small molecule pharmaceuticals at the time of assay validation **Proposed change :** Replace" fully characterized standards" with " fully characterized to the extent possible**"**	'Fully' has been changed into 'sufficiently' for clearness.
108	23	"Suitability of the reference standard should be scientifically justified". Does that only concern the fact to have a Certificate of Analysis or is additional scientific rationale expected? It is said "The use of certified standards is not needed	The text 'Suitability of the reference standard should be scientifically justified' has been deleted as this is considered obvious from the previous text.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		for IS, as long as…," but in line 114 stated that it is essential that the labelled standard is of the highest isotope purity… The text should clarify if there is the necessity for a certificate of analysis for labelled standards used as IS or identify that the text is speaking about standards of reference only.	It is considered that for an internal standard a Certificate of Analysis is not necessary, as long as the suitability for use as internal standard has been proven.
108-109	19	What do you mean on "Suitability of the reference standard should be scientifically justified"? It is not clear from the text.	The text 'Suitability of the reference standard should be scientifically justified' has been deleted as this is considered obvious from the previous text.
108-109	22	Comment: Point 4.1 Reference standards Clarification is sought as to what is meant by "Suitability of the reference standard should be scientifically justified".	The text 'Suitability of the reference standard should be scientifically justified' has been deleted as this is considered obvious from the previous text.
108-113	47	Comment: Regarding the statement that "Suitability of the reference standard should be scientifically justified." This is logical, and the suitability of the specific reference standard should be documented with a Certificate of Analysis (COA) and/or supporting experimental data to address interference, stability in matrix, and reconstitution solution or assay and stop buffers, etc. The next sentence, "The use of certified standards is not needed for IS, as long as the suitability for use is demonstrated," requires clarification since the stability, percentage of isotopic labeling, purity, extraction tracking, stability in reconstitution solutions and buffers relative to the reference standard or solubility, to name a few, are critical to the ruggedness of the method and quality and integrity of the data, and therefore requires the same degree of documentation and/or experimentally demonstrated suitability. The sentence as it reads could allow a laboratory to arbitrarily choose an internal standard based on convenience without adequately demonstrating suitability and subsequent effect on method ruggedness.	The text 'Suitability of the reference standard should be scientifically justified' has been deleted as this si considered obvious from the previous text. It is considered that for an internal standard a Certificate of Analysis is not necessary, as long as the suitability for use as internal standard has been proven.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Many of the incurred sample analysis failures observed in industry are a result of reference standard or internal standard stability issues, or the use of an inappropriate internal standard **Proposed:** Recommend that prescribing adequate documentation, such as a COA and/or experimental acceptability for the internal standard that is determined during method development, and demonstrated and documented in method validation, will ensure the use of suitable internal standards. Clarification within the guidance for the inclusion of an adequate COA and/or respective method validation documentation will ensure that respective lots of the internal standard used during sample analysis are appropriate and stability, isotopic labeling, ionization, etc. are adequate for the validated method.	
109 - 111	20	Comment: I agree with the provided text, but IS stock solution stability should also be addressed to ensure suitability of use over time **Proposed change (if any):** Propose including IS stock solution stability	Stock solution stability is covered under section 4.1.9 Stability of this guideline.
109-111	48	Comment: The analysis certificate for internal standard is not request by this draft guidance neither in the study file nor in the validation report provided that its suitability is demonstrated for the study. This point means that no deviation to the GLP compliance has to be recorded in case of lack of analysis certificate or expiration/retest date relating to the internal standard.	Comment noted.
109-113	37	Comment: The suitability of IS is of general knowledge, moreover it is discussed in other paragraphs. What does "e.g. lack of interference is shown for the substance itself or any impurities thereof" mean? Are there any additional experiments required for impurities	Lack of interference has been changed into 'lack of analytical interference'. The proposed text is not agreed. It is considered that for an internal standard a Certificate of Analysis is not necessary, as

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		of IS? If yes they should be clearly defined. **Proposed change (if any):** The use of certified standards is not needed for IS, ~~as long as the suitability for use is demonstrated, e.g. lack of interference is shown for the substance itself or any impurities thereof. Whoever the supplier,~~ but a certificate of analysis is required to ~~ensure~~ define manufacturer ~~quality, stability,~~ storage conditions, expiration date, batch number and purity of the reference standards.	long as the suitability for use as internal standard has been proven.
110	36	**Comments** Clarification in the text should be given as to the need for certificates of analysis for internal standard and whether they are needed or not.	It is considered that for an internal standard a Certificate of Analysis is not necessary, as long as the suitability for use as internal standard has been proven. This is clearly stated in the revised text.
112 - 113	3	Required documentation should include lot number, concentration, stability and the *monitoring of expiration dating*.	The text has been changed and is considered sufficient clerar. For LBA, reference standards being mostly solution, concentration is of importance. A separate section has been introduced on ligand binding assays.
112-113	8	Comment: Certificate of analysis needed to ensure quality of the reference standard. In previous paragraph (line 103), quality was already defined as purity, which is also stated later on line 113. **Proposed change (if any):** Remove the word "quality" since purity is already mentioned.	The text has been revised and quality has been deleted. In addition, the same applies to stability.
112-113	45	1. CoA for method validation not necessarily needed, the test item should be of suitable identity and purity. For study conduct availability of a CoA is a must. 2. For IS purity and CoA are not needed.	It is considered that for an internal standard a Certificate of Analysis is not necessary, as long as the suitability for use as internal standard has been proven. This is clearly stated in the revised text.
114-117	23	A wide range of types analytes are used by industry from simple organic molecules over proteins and nucleic acids to microorganisms. **Proposed change (if any):** Due to the large number of different analytes the examples given should be	The comment is not clear. Furthermore, microorganisms and nucleic acids are not within the scope of this guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		reviewed and improved.	
114-117	37	Comment: It seems to be an overregulation. Most of bioanalytical methods published in literature use as IS another substance rather than labelled analyte. Those methods were validated and produce reliable results. Moreover, the labelled IS is not always a proper choice (e.g.: J. Wieling, Chromatographia 55 Supp. (2002) S107-S113). Generally, the labelled standards are more expensive than un-labelled. To prevent different interpretations in different EU member states, it should be clearly stated that another substance than labelled analogue used as IS is an acceptable approach. **Proposed change (if any):** When MS detection is used in the bioanalytical method, a stable isotope-labelled IS ~~is recommended to~~ may be used ~~whenever possible~~ if justified.	As indicated, it is a recommendation and not a requirement.
116	6	Comment: **Proposed change (if any):** Replace the last sentence with new sentence: A correction factor for the contribution of the unlabelled analyte in isotope-labelled IS to the total peak-area of IS should be used.	The text has been revised and considered sufficiently clear now.
116	9	The internal standard should be of adequate purity for the intended use, it may not be necessary to have the "highest" purity. **Proposed change (if any):** The statement that it is "essential that the labelled standard is of the highest isotope purity" should be reworded.	The text has been changed into ' However, it is essential that the labelled standard is of the highest isotope purity and that no isotope exchange reaction occurs. The presence of any unlabelled analyte should be checked and if relative amounts of unlabelled analyte are detected the potential influence has to be evaluated during method validation.'
116	18	Comment: **Proposed change (if any):** Replace the last sentence with new sentence: A correction factor for the contribution of the unlabelled analyte in isotope-labelled IS to the total peak-area of	The text has been changed into ' However, it is essential that the labelled standard is of the highest isotope purity and that no isotope exchange reaction occurs. The presence of any unlabelled analyte should be checked and if relative amounts of unlabelled analyte are detected the potential influence has

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		IS should be used.	to be evaluated during method validation.’
116-117	14	Comment: The presence of any unlabelled analyte in isotope-labelled internal standard material will not generally bias the results. It rather influences the LLOQ of the method. **Proposed change:** Please delete the sentence or replace the sentence by: *“If relevant amounts of unlabelled analyte are detected the potential influence has to be evaluated during method validation.”*	The text has been changed into ‘ However, it is essential that the labelled standard is of the highest isotope purity and that no isotope exchange reaction occurs. The presence of any unlabelled analyte should be checked and if relative amounts of unlabelled analyte are detected the potential influence has to be evaluated during method validation.’
116-117	21	It should be recognized that unlabeled analyte will always be present to some extent in a stable labeled IS. **Proposed change (if any):** We would recommend to add: *The amount of unlabeled analyte allowed to be present without introducing unacceptable bias should be defined in SOPs*.	The text has been changed into ‘ However, it is essential that the labelled standard is of the highest isotope purity and that no isotope exchange reaction occurs. The presence of any unlabelled analyte should be checked and if relative amounts of unlabelled analyte are detected the potential influence has to be evaluated during method validation.’
116-117	39	Of course, it is preferable to use an isotope-labelled IS which does not contain any unlabelled analyte. But if this is not possible, the software of most LC-MS systems enables to use a correction factor for the contribution of unlabelled analyte in IS to the total peak area. **Proposed change (if any):** Add a new sentence: In such a case a correction factor for the contribution of unlabelled analyte in IS to the total peak area must be used in the quantitation software.	The text has been changed into ‘ However, it is essential that the labelled standard is of the highest isotope purity and that no isotope exchange reaction occurs. The presence of any unlabelled analyte should be checked and if relative amounts of unlabelled analyte are detected the potential influence has to be evaluated during method validation.’
118 - 123	24	Comment: For chromatography based assays some companies would prefer to test less then 6 animal matrix sources as the matrices are more standardized as they are for humans. Could we ask the EMA to comment on the measurement of endogenous compounds? Please specify the requirements for LBA here as well.	It is agreed that there may be less variability but this does not exclude testing in a sufficient number. In case of endogenous compounds, a matrix can be selected not containing this compound, for instance after a clean-up. Still, the matrix should be representative for the matrix of study samples. Validation of an analytical method may be subject to a case by case evaluation.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
118-139	2	Comment: • The amount of 6 sources to be tested is suitable for human matrix validations. However, in view of possible availability issues and known uniformity for animal sources, a number of at least 2 sources would be more appropriate and largely sufficient in case of animal matrix validations. • Co-medication: In clinical studies with only one co-medication or in DDI clinical studies it makes sense. Otherwise it will be overkill if all the co-medications occurring in the study need to be addressed. • We agree that it is important to demonstrate the back-conversion during the course of analysis. However, there should be no need to demonstrate it for each step in the analysis. Also a determination of total analyte should be sufficient to provide scientific evidence for the need to change assay practices to reduce or eliminate back-conversion. • The requirement to spike samples with the metabolite at the highest in vivo concentrations is not always possible as the metabolites, unless of significance as defined by the rules for safety testing, are not routinely synthesized. • For large molecules typically ≥10 blank matrices are tested. • For LBAs it is also important to show recovery of	It is agreed that there may be less variability but this does not exclude testing in a sufficient number. This is agreed. The text has been revised for clarity by including 'on a study specific and compound specific base'. Evaluation should cover the whole procedure, this includes start and end. The procedure how to evaluate back-conversion has been deleted. This is considered too detailed for a guideline. Furthermore, there may be other methods for evaluating this issue. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		LLOQ levels of spiked drug in context of these matrices – simply demonstrating blank levels alone is insufficient. This requirement is in conflict with lines 334-336 that stated blank response may exceed 20% … as long as it does not affect accuracy. • For ligand binding assays, selectivity is tested by analyzing blank matrix from (at least 10) different sources (different batches, individual sources including disease matrix, when applicable) and the same matrix lots spiked with the analyte at a defined concentration, e.g. at low QC level. Absence of interfering components is accepted where the response is below the lower limit of quantitation and the recovery of the spiked analyte is within the criteria defined for the low QC. It may be prudent also to evaluate selectivity at higher analyte concentrations, as well. • Due to interference of endogenous compounds, the blank response may exceed 20% of the LLOQ; however, this may be acceptable, as long as it does not affect accuracy. **Proposed change (if any):** • This section describes the procedures of selectivity assessment for chromatographic assays. It is recommended to have this section within the specific section of chromatographic assays and describe the selectivity assessment for LBA in the proposed separate section.	assays.
118-139	21	We consider that evaluation of back conversion of a metabolite is a method development activity and shouldn't be included in a regulatory guidance.	This may be an issue for method development, however this issue is considered important and therefore it is included in this guideline. The procedure how to evaluate back-conversion has been deleted. This is considered too detailed for a guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
118 -139	36	**4.1.1 Selectivity** Greater clarity and guidance is requested for this section and the proposed validation requirements. In particular the proposed requirements on the extent of expected validation with respect to co-medications , the investigations required to revalidate methods as "further knowledge" of metabolism is gained and the investigations on "interference from degradation products formed during sample analysis".	The text has been revised for clarity.
118-139	45	It provides a nice experiment to evaluate back conversion of a metabolite, but it is too specific and is a method development like activity to be included in a regulatory guidance.	This may be an issue for method development, however this issue is considered important and therefore it is included in this guideline. The procedure how to evaluate back-conversion has been deleted. This is considered too detailed for a guideline.
118 – 256	13	Comment: The guideline focuses on chromatographic methods and only marginally addresses other assays, which are key to our studies. Because the chromatographic methods are covered first, this suggests that everything which applies for this type of method also applies for the ligand-binding assays, which are addressed later. We would recommend a separation of these different types of assays and to define separate requirements wherever appropriate (see general remarks) **Proposed change (if any):** We would propose to clearly separate these different types of assays and to define separate requirements wherever appropriate. The ligand binding assays do not cover all classes of biological assays in use to date. There should also be a focus on biological assays described, e.g. in the EP, Chapter 2.7, in which the samples are tested in dilution series (see general remarks).	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
118-256	52	Comment: The guideline focuses on chromatographic methods and only marginally addresses other assays relevant to the medicinal products manufactured by	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		PPTA members. Because the chromatographic methods are covered first, this suggests that everything which applies for this type of method also applies for the ligand-binding assays, which are addressed later. **Proposed change (if any):** We would propose to clearly separate these different types of assays and to define separate requirements wherever appropriate. The ligand-binding assays do not cover all classes of biological assays in use to date. There should also be a focus on biological assays described, e.g. in the EP, Chapter 2.7, in which the samples are tested in dilution series (see general remarks).	separate section has been introduced on ligand binding assays.
119-123	19	There is no criterion for internal standard, however it is important to prove the selectivity for IS, too. **Proposed change (if any):** We suggest to change the text as follows: "Absence of interfering components is accepted where the response is less than 20% of the lower limit of quantitation for the analyte and 5% for the internal standard."	The text has been changed into 'Normally, absence of interfering components is accepted where the response is less than 20% of the lower limit of quantification for the analyte and 5% for the internal standard.'
119-123	21	Within this section, some of the proposed requirements are consistent with internationally accepted standards, e.g. this section states that 6 lots of matrix should be tested, while others are not, e.g. this draft guideline does not allow for any lots to contain interferences. **Proposed change (if any):** We would suggest that the final guideline should allow one lot to show interferences.	The guideline does not state that this is not allowed. If a positive interference is observed in 1 sample, the applicant takes the risk to expect interference in the study samples.
119-123	45	1. Interference should refer to a peak or signal occurring at the RT of the analyte in the blank matrix sample (chromatography). 20% acceptance is valid for chromatography based assays. For ELISA the assay must be able to distinguish between the LLOQ and the blank signal based on	1. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		the statistics of at least 10 observations of individual blank matrix samples. 2. Use of specific mass spectrometry and stable isotope labeled IS makes check of 6 sources obsolete. For ELISA the number of individual sources of matrices to be tested for selectivity should be at least 10.	2. This comment is unclear as regards how LC-MS can compensate for interference.
119-139	21	Comment: Point 4.1.1 Selectivity The possibility of back-conversion of metabolic species may be proved during the validation if: 1. the metabolism is known 2. the reference standards are commercially available 3. the reference standards are of sufficient purity The lack of back-conversion may be evidenced by consideration of the reproducibility of study samples (ISR) in different stability tests (e.g. short term, freeze and thaw, post preparative, ...) or by using the incurred sample analysis to rule out back-conversion as a possibility. Testing of back-conversion of metabolites as currently suggested in the draft guideline is usually not feasible due to non-availability of the metabolites for testing. The EGA would strongly recommend that alternative approaches should therefore be acceptable for the reanalyzed sample, (e.g. Incurred Sample Reanalysis), possibly in combination with certain storage periods (for which stability of the parent compound has been demonstrated) at room temperature.	The text has been revised for clarity. The procedure how to evaluate back-conversion has been deleted. This is considered too detailed for a guideline.
120-121	50	Comment: Selectivity should be proven by using at least 6 sources of the appropriate blank matrix Further clarification on acceptance criteria for selectivity	The guideline does not state that this is not allowed. If a positive interference is observed in 1 sample, the applicant takes the risk to expect interference in the study samples.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		is required e.g. "absence of interfering components is accepted where the response is less than 20% of the lower limit of quantitation for the analyte", here % of accepted lots (80% or 100%; 6 out of 6 or 6 out of 10 lots) should also be incorporated to provide the required clarity and understanding.	
121	9	**Proposed change (if any):** A statement that fewer sources may be acceptable for rare matrices should be included.	The text has been revised accordingly.
121	19	What do you mean on "6 sources of the appropriate blank matrix"? Does it mean matrices from 6 individual subjects or animals, or pools of 6 different sub-populations? **Proposed change (if any):** Please include examples.	The text has been revised and states now "6 individual sources'.
121	20	Comment: The 6 sources of plasma used for selectivity and matrix effect evaluation should be obtained from individual donors (or animals) rather than lots pooled from multiple donors or sources **Proposed change (if any):** Recommend clarifying that the plasma lots must be from individual donors or animals	The text has been revised and states now "6 individual sources'.
121	22	Comment: Point 4.1.2 Selectivity It is unclear what is meant by "6 sources of the appropriate blank matrix". Does it refer to matrices from 6 individual subjects or to pools of 6 different sub-populations?	The text has been revised and states now "6 individual sources'.
122-123	37	Comment: This sentence is not clear. There are no acceptance criteria for IS. **Proposed change (if any):** It is accepted that selectivity is ensured if: a) For analyte: response (peak area or height) in blank sample is less than 20% of response to analyte in LLOQ	The text has been changed into 'Normally, absence of interfering components is accepted where the response is less than 20% of the lower limit of quantification for the analyte and 5% for the internal standard.'

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		sample; b) For IS: response (peak area or height) in blank sample is less than 5% of response to IS in zero sample.	
122-123	40	Comment: There are no criteria discussed regarding the level of interference for the Internal Standard (IS). **Proposed change (if any):** Add an acceptance criteria for interference at the IS retention time (i.e. industry standard is 5% of the mean response of the IS of all acceptable standards).	The text has been changed into 'Normally, absence of interfering components is accepted where the response is less than 20% of the lower limit of quantification for the analyte and 5% for the internal standard.'
123	8	Comment: Unclear if acceptance criterion refers only to analytes of interest, or IS as well. If not for IS, there should be a criteria for the IS as well.	The text has been changed into 'Normally, absence of interfering components is accepted where the response is less than 20% of the lower limit of quantification for the analyte and 5% for the internal standard.'
123	11	Comment: add "that of" after "20% of"	The text has been changed into 'Normally, absence of interfering components is accepted where the response is less than 20% of the lower limit of quantification for the analyte and 5% for the internal standard.'
123	23	What about the selectivity regarding endogenous compounds? Could this test not be done? The % of the response of interfering components does not seem to be established for IS. Given that there is an acceptability range for the analyte is it necessary to apply another for the internal standard? If yes, what should this be and on what basis should it be chosen?	In case of endogenous compounds, a matrix can be selected not containing this compound, for instance by a clean-up of the matrix. Still, the matrix should be representative for the matrix of study samples. Validation of an analytical method may be subject to a case by case evaluation. The text has been changed into 'Normally, absence of interfering components is accepted where the response is less than 20% of the lower limit of quantification for the analyte and 5% for the internal standard.' This may be possible, but this is subject to further method development evaluation. Use of another IS will for instance impact on the analyte/IS ratio and calibration response function.
124	23	The metabolites investigation may present technical difficulties if the metabolites are not available in the shops (feasibility to have study samples before the	This is acknowledged. The text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		validation). Same for the "back-conversion".	
124-127	7	Comment: The expression "co-medication" should focus on other medications administered in the study but should not include "allowed medication" taken by the subject At the time of method validation one doesn't know all co-medications that may be used later on. **Proposed change (if any):** Please consider re-phrasing, e.g. *"Co-medications normally used in the subject population studied should be taken into account at the stage of method validation, or on a study specific base."*	The text has been revised accordingly to the proposed text.
124-127	21	A clear definition of the meaning of "co-medication" in the context of method validation should be introduced. Checking for interference with co-medication is not necessarily part of method validation but should be included in the in-study validation for DDI studies. At the time of method validation it is almost impossible to predict all co-medications that may be used later on during development. Interference from co-medication should only be investigated when justified, i.e. when medication with structural similarities to the investigative drug is present. Otherwise there would potentially be too many potential drugs to test for interference. **Proposed change (if any):** Therefore we would recommend to indicate that e.g. Co-medications normally used in the subject population studied should be taken into account at the stage of method validation, or on a study specific and compound basis.	The text has been revised accordingly to the proposed text
124 - 127	24	Comment: We recommend the tiered approach for this section.	Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			subject to case by case evaluation.
124-127	37	Comment: This point is not clear. More details are needed: a) when each investigation is required? b) how it should be performed? c) what are acceptance criteria? Metabolites: a) should all possible metabolites be investigated or only over certain limit (e.g. 30% of parent compound)? b) How to asses this limit if it is required? Is literature data acceptable? Or is metabolite identification study necessary to perform method validation for a BE study? Interference from degradation products formed during sample preparation: a) What is the reason for it? Both calibration/QC and study samples are processed so degradation during sample preparation should not influence the measurement result. b) How to compare processed sample with sample without sample preparation? **Proposed change (if any):** It may also be necessary to investigate the extent of any interference caused by major metabolites of the drug(s) (reaching over XX% of parent drug concentration), ~~interference from degradation products formed during sample preparation~~, and interference from possible co-administered medications (co-medications normally used in the subject population studied). ~~Co-medications normally used in the subject population studied should be taken into account.~~	This is considered too detailed for this guideline. This is considered too detailed for this guideline. The proposed text is not agreed. Instead the text has been revised as follows for clarity: 'It may also be necessary to investigate the extent of any interference caused by metabolites of the drug(s), interference from degradation products formed during sample preparation, and interference from possible co-administered medications. Co-medications normally used in the subject population studied which may potentially interfere should be taken into account at the stage of method validation, or on a study specific and compound specific base.'
124-127	45	1. Check for co-medication interference is not necessarily a part of the validation. But, should	The text has been revised as follows for clarity:

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		be included in the in-study validation for DDI studies (both chromatographic and biological methods). 2. Check for co-medication might be impossible for all possible co-meds, in DDI warranted, in case of LC-MS interference is unlikely due to mass selective detection. 3. Clarify that testing only co-meds that could potentially interfere. (Particularly important because in Ligand Binding Assays, co-meds which are small molecules are not anticipated to interfere). **Proposed change (if any):** Co-medications with the potential to interfere in the method should be tested for interference	'It may also be necessary to investigate the extent of any interference caused by metabolites of the drug(s), interference from degradation products formed during sample preparation, and interference from possible co-administered medications. Co-medications normally used in the subject population studied which may potentially interfere should be taken into account at the stage of method validation, or on a study specific and compound specific base.'
124-139	23	A precision could be included to identify which metabolites are to be followed and how (e.g. principle metabolite, % circulating level, list of potentially unstable metabolites etc.). It may also be necessary to investigate: the extent of any interference caused by metabolites of the drug(s), interference from degradation products, etc. The text should explain the procedure to investigate the interferences. On the other hand, the text explains that the possibility of back-conversion of a metabolite into parent analyte should be evaluated but does not explain how and when (in pre-clinical and/or clinical part?)	This is considered too detailed for this guideline. The text has been revised as follows for clarity: 'Co-medications normally used in the subject population studied which may potentially interfere should be taken into account at the stage of method validation, or on a study specific and compound specific base.'
125	21	We do believe that it should not be necessary to investigate "interference from degradation products formed during sample preparation". Any possible interference will become apparent during the regular course of validation. Therefore specific investigation should not be necessary.	The guideline does not state 'should' but indicate that 'it may be necessary'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
126	26	A thorough evaluation of co-medications might be difficult.	This is agreed. The revised text indicates now 'at the stage of method validation, or on a study specific and compound specific base.'.
126-127	31	Comment: The guideline states that co-medications normally used in the subject population study should be taken into account for investigation of interferences. In case of phase III studies in which pharmacokinetic assessments are done, it may be impossible to investigate all co-medications. Depending on the population, e.g. chronic kidney disease patients, other chronic diseases, patients take a number of different co-medications. **Proposed change:** … may be taken into account for special populations (e.g. patients taking several co-medications).	The revised text indicates now 'at the stage of method validation, or on a study specific and compound specific base.'
127 - 139	24	Comment: This paragraph is not in the FDA guidance. Addition of this text is welcomed in general as it adds clarity. However, it should be recognized that metabolite standards are sometimes very difficult to synthesise, e.g. phase II metabolites like glucuronides – and if they are not active (which is usually the case with glucuronides) then as long as their back-conversion potential is appropriately controlled (e.g. acidification) and they are chromatographically separated from the analyte of interest, there should be no problem. We recommend the tiered approach for this section.	The problem is recognized. The text has been revised for clarity. Tiered approaches may be acceptable when sufficiently supported, however acceptability of tiered approaches will be subject to case by case evaluation.
128-139	1	The proposed approach to evaluate back-conversion only with blank matrix samples spiked with synthetic metabolite is not so realistic due to the following reasons: A) The metabolic transformation of many well known registered molecules is quite poor B) The availability of synthetic metabolites is quite poor especially of those chemically unstable C) The costs involved in producing and/or purchasing synthetic metabolites can be too high	The text has been revised and includes the possibility to use study samples.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change: Back-conversion has to be tested with incurred samples of volunteers and/or patients treated with the drug under study comparing the quantitative results in different conditions of processing and storage relevant for the method under development. **Remark:** It is important to note that the incurred samples reanalysis procedure, as presented in page 13 lines 468-486, aiming to check the analytical method reproducibility in each particular study, should remain as proposed. The back-conversion tests on incurred samples would be only a validation procedure to detect this kind of analytical interference.	Comment noted. The section on incurred sample reanalysis is kept in this guideline.
128-139	6	Comment: This is correct for **unstable** metabolites. But with this wording we are afraid that this testing will be required e.g. for all glucuronide metabolites, majority of which are stable compounds. Here the importance of literature data should be emphasized.	The text has been revised and 'unstable' has been included.
128-139	7	Comment: Comment on entire paragraph: For labile metabolites it is very hard to obtain a reference material even by "bio-synthesis". To demonstrate sample stability life sample stability should be considered an appropriate tool.	The text has been revised and includes the possibility to use study samples.
128-139	14	Comment: It is necessary that the influence of instable metabolite-(conjugates) on the parent analyte should be investigated in the case of hints for the presence of such metabolites. However, this can only be performed by spiking the potentially unstable metabolite-(conjugate, N-oxides) to the matrix. If the conjugate cannot be easily assessed (which is very often the case) a reanalysis of a stressed sample is recommended in order to evaluate a potential increase of signal of the parent drug. The latter investigation will be performed after	The text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		availability of study samples during incurred samples reanalysis. **Proposed change:** Please add a section what to do in case of instable metabolite-(conjugates) which cannot be spiked to the matrix since they are not available as reference material.	
128 - 139	16	Comment: We agree that it is important to demonstrate the back-conversion during the course of analysis. However, there should be no need to demonstrate it for each step in the analysis; a determination of total should be sufficient to provide scientific evidence for the need to change assay practices to reduce or eliminate back-conversion. Additionally, the requirement to spike samples with the metabolite at the highest in vivo concentrations is not always possible as the metabolites, unless of significance as defined by the rules for safety testing, are not routinely synthesized. **Proposed change (if any):** Please allow the laboratory to define scientific approaches that assess with confidence the absence of a back-conversion during the analysis.	The procedure how to evaluate back-conversion has been deleted. This is considered too detailed for a guideline**.**
128-139	19	The issue of back-conversion of metabolites into parent drug has got higher importance in the guideline than it has in practice. Performing this test is difficult since unstable metabolites are often not available. The issue can be covered by incurred sample reanalysis. **Proposed change (if any):** We suggest to omit this study unless back-conversion is decribed in the literature.	The text has been revised and includes the possibility to use study samples.
128 - 139	20	Comment: Some glucuronides have also been shown to undergo conversion from metabolite to parent in the MS source. If parent and glucuronide metabolites co-elute (as with a ballistic gradient), this conversion may lead to a positive bias in determined parent concentration in	The text 'or in the MS ion source' has been added.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		actual samples **Proposed change (if any):** Suggest evaluating the potential for in-source conversion of labile metabolites to parent and recommend as a general approach, chromatographic resolution of parent and metabolites to minimize this possibility	
128-139	21	For labile metabolites it is very hard to obtain a reference material even by "bio-synthesis". We would, therefore, recommend that to demonstrate sample stability, life sample stability should be appropriate. With reference to the MIST guidance we would also recommend that checking for metabolite back-conversion should only apply to known metabolites and/or when >10% of exposure compared to the entire drug material.	The text has been revised and includes the possibility to use study samples. Adding a cut off value is considered too detailed for this guideline.
128-139	39	This is correct for **unstable** metabolites. But with this wording we are afraid that this testing will be required e.g. for all glucuronide metabolites, majority of which are stable compounds. Here the importance of literature data should be emphasized.	The text has been revised for clarity and indicates now also 'unstable metabolites'.
128-139	45	1. Check for metabolite back-conversion only for known metabolites and/or >10% of exposure compared to entire drug material (see MIST guidance) 2. It should be assessed during validation if stable isotope labeled IS contains unlabeled material which causes a bias. The level of unlabeled material may not bias the quantification of analyte, if the concentration of IS is sufficiently low	1. Adding a cut off value is considered too detailed for this guideline. 2. The issue of unlabelled material is already indicated in the Reference standards section under paragraph 4.1.
131-135	8	Comment: The test described is very specific. Leeway should be given for equivalent test designs.	The text has been rewritten for clarity. The procedure how to evaluate back-conversion has been deleted. This is considered too detailed for a guideline.
133	8	Comment: "Actual" highest in vivo concentrations are	The text has been rewritten for clarity. The procedure how to

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		often not known, or the literature is not correct. **Proposed change (if any):** Change "actual" to read "expected"	evaluate back-conversion has been deleted. This is considered too detailed for a guideline.
134	36	**Comments** Clarity on how the "chromatograms are to be evaluated for the formation of parent analyte " in a selective assay such as LC-MS/MS is requested.	The text has been rewritten for clarity. The procedure how to evaluate back-conversion has been deleted. This is considered too detailed for a guideline.
140	34	Comment: Carry-over is specific to HPLC method. **Proposed change (if any):** Specify in the text that carry-over should be assessed for non ligand-binding assays	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
140	35	Comment: The carryover as defined "is the appearance of an analyte signal in blank sample peaks after the analysis of samples with a high analyte concentration". The method where the response in blank sample is less than 20 % of the LLOQ (acceptable selectivity) may not be accepted for carryover. Thus the limits for carry over need to be defined. **Proposed change (if any):** The limit could be less than 20 % of the analyte response observed in LLOQ and 5 % response of that IS response observed in LLOQ.	The text has been revised and a limit of 20 % of the LLOQ and 5 % of the IS has been introduced.
140-143	50	Comment: Carry over should be assessed by injecting blank samples after a high concentration sample or calibration standard. Here, it is useful to mention the exact concentration high or highest on calibration curve. A Carry-over acceptance criterion is missing in carry-over section. Appropriate acceptance criteria for carry –over should be incorporated to this guidance in carry-over section to avoid any bias later on. For example, carry-over blank	The text has been revised and 'ULOQ' has been included in the text. The text has been revised and a limit of 20 % of the LLOQ and 5 % of the IS has been introduced.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		followed by injection of highest standard on calibration curve should not be more than 20% of the LLOQ response.	
140-149	2	Comment: • Randomization should still be possible if the carry-over is shown to be sufficiently low to allow randomized set-up of the run. • Randomization can be very helpful in LBA's to reduce bias from a plate effect. **Proposed change (if any):** • Clearly state: for chromatographic methods or for automated ligand binding assays carry-over should be addressed….OR … Carry-over should be addressed and minimized during method development where applicable/ as appropriate, e.g. for chromatographic methods or automated ligand binding assays. • Exclude LBA's from this recommendation for carry-over in chromatographic-based methods	The text regarding avoiding randomisation has been revised for clarity. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
140 - 149	24	Comment: Please include "during chromatographic analysis" for carry-over. Delete randomization for validation. **Proposed change (if any):** Should the title include cross contamination?	The text regarding avoiding randomisation has been revised for clarity.
140-149	45	1. For small molecules (chromatography methods), suggest assessing carryover by injecting blank samples after ULOQ. ULOQ represents the worst case scenario. Also suggest having , for example: >20% of LLOQ. 2. Not applicable for ELISA, only in case robots are used for sample prep (with fixed needles	The text has been revised and 'ULOQ' has been included in the text. The text has been revised and a limit of 20 % of the LLOQ and 5 % of the IS has been introduced. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			separate section has been introduced on ligand binding assays.
141	8	Comment: The use of "may not" at the end of the line is not specific enough. **Proposed change (if any):** Change "may not" to read "must not"	The text has been deleted for clarity.
141-142	37	Comment: This point is inconsistent with selectivity acceptance criteria, e.g. in case of the method with range 1-100 ng/ml the highest acceptable interference during selectivity assessment is ca. 0,2 ng/ml while highest carry over accepted is 100 x 15% = ca. 15 ng/ml. The carry-over acceptance criteria should be consistent with selectivity acceptance criteria. **Proposed change (if any):** Carry-over may not affect method selectivity (see section 4.1.1). ~~accuracy and precision (see section 4.1.5 and 4.1.6).~~	This text has been deleted for clarity.
141-147	19	There are no criteria and procedure for performing carry-over test. **Proposed change (if any):** We suggest to perform this test by replicates (e.g. injecting blank samples after high concentration samples six times) and use the acceptance limits of selectivity (20% for the analyte and 5% for IS) for at least 67% of replicates.	The text has been revised and a limit of 20 % of the LLOQ and 5 % of the IS has been introduced.
141-147	40	Comment: The guideline proposes to assess carry-over in the validation stage. However, carry-over seen at the validation stage does not mean you would see it again at the time of sample analysis. The reverse is true due to the differences in instrument sensitivity, especially when multiple instruments are involved in the laboratory. I believe it would be important to test the carry-over before initiating the sample analysis with the instrument you are planning to use for the study sample analysis. This could be done as part of system suitability for example.	The proposal is not deemed appropriate for a guideline but should be addressed in a SOP.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): Regardless of the outcome, the carry-over should also be tested before initiating the sample analysis.	
142	36	**Comments.** No guidance is offered as to what criteria for acceptability are to be used. Is this based on signal relative to the LLOQ sample or is it left to sponsor derived criteria	The text has been revised and a limit of 20 % of the LLOQ and 5 % of the IS has been introduced.
143	8	Comment: Carryover tests involving a "high" standard is not specific enough. **Proposed change (if any):** Change "high standard" to "highest standard"	The text has been revised and 'ULOQ' has been included in the text.
143	20	Comment: All efforts possible should be expended to eliminate carry-over during method development (MD). These experiments should be based on the analyte chemistry and should include parameters such as pH adjustment, organic solvent type and strength, wash solvent volume, etc. Only when all reasonable options have failed should the specific measures in the next sentence be considered **Proposed change (if any):** Recommend adding a new sentence between the two current sentences stating that all efforts to eliminate or minimize carry-over should be made, and that the specific measures in the next sentence should be implemented only as a last resort	It is considered that the text is sufficient clear.
143	49	Blank injections - is this a matrix blank or a mobile phase/recon solution?	During validation matrix blank; during study samples analysis in case carry over is observed, blank matrix or mobile phase/reconstitution solution.
143-144	23	The text states "If it appears that carry-over is unavoidable, specific measures should be considered…." The text should define what would be these specific	The proposal is not deemed appropriate for a guideline but should be addressed in a SOP.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		measures. In this section there are no acceptability levels. Is it not possible to attach these to other acceptance limits such as those for selectivity?	The text has been revised and a limit of 20 % of the LLOQ and 5 % of the IS has been introduced.
145	26	If blanks are used to address carryover, they should be inserted between each injection for quantitation. **Proposed change (if any):** Change text accordingly	This depends on expected concentrations and at which concentration carry over occurs.
147	20	Comment: If specific measures (lines 144 – 147) are implemented to deal with carry-over, a clear process should be included in the method for dealing with samples with concentrations > ULOQ. Procedures for determining the extent of bias in the following samples and criteria to determine which samples are impacted to the extent that reanalysis will be required should be included **Proposed change (if any):** Suggest that a specific plan to address the extent of carry-over and impact in subsequent samples should be included in the method when specific measures to address carry-over are implemented	The proposal is not deemed appropriate for a guideline but should be addressed in a SOP.
148-149	8	Comment: If the absence of carryover is demonstrated, then it should not be mandated that randomisation be avoided.	The text has been revised for clarity.
148 - 149	20	Comment: One way to ensure an accurate calibration curve (unaffected by carry-over) is included in each batch is to inject a calibration curve at the beginning of each batch, blanks and zero standards first, then calibration standards injected from low to high concentration. If carry-over is present in a batch, samples and QCs can be compared to this initial curve to determine the extent of any bias present **Proposed change (if any):** Consider including this as a possible recommendation at the end of line 149	This only applies to the calibration curve, but the problem still may excist for the study samples.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
148-149	30	On the contrary, randomisation should be used here as well! The carry over can then be understood from residual analysis, plotting Δc = ci-ci-1 against ri=yi-ymean.	The comment is unclear.
150	47	Comment: Regarding the discussion of the lower limit of quantitation (LLOQ). In some cases, the LLOQ can also be the next highest standard should the validated LLOQ standard not pass acceptance criteria within a specific run and the method allows for calibration curve truncation. **Proposed:** Recommend that additional discussion be included to discuss the appropriate use of LLOQs when a calibration curve is truncated in a run.	This is included in section 5.2 Acceptance criteria of an analytical run.
150-153	21	**Proposed change (if any):** The lower limit of quantitation (LLOQ) is the lowest amount of analyte in a sample, which can be quantified reliably, with an acceptable accuracy and precision (see Accuracy and Precision). *In addition, for chromatographic assays, the analyte signal of the LLOQ sample should be 5 times the analyte signal of a blank sample.* The LLOQ should be adapted to expected concentrations and to the aim of the study.	The text has been (partial) revised to include the 5 times limit.
150 - 153	24	Comment: We would like to ask EMA to align with FDA definition of LLOQ.	The EMA considers harmonisation an important issue. Harmonisation issues have been taken into account as much as possible.
150-153	45	In addition, for LC-MS based assays the analyte signal of the LLOQ sample should be 5 times the analyte signal of a blank sample. No interference peak greater 20% of LLOQ at RT of analyte should be present	The text has been (partial) revised to include the 5 times limit.
151	20	Comment: In practice, an assay LLOQ cannot be defined at an analyte concentration lower than the lowest acceptable calibration standard used during method validation	The text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): Suggest revising this sentence to state that the assay LLOQ will be defined as the lowest calibration standard in the method with acceptable accuracy and precision determined during assay validation	
151-153	37	Comment: According to FDA minimal signal to noise ratio is required (S/N ≥ 5). **Proposed change (if any):** The signal to noise ratio of minimum 5 is required for LLOQ (S/N ≥ 5).	The text has been (partial) revised to include the 5 times limit.
151-153	43	The lower limit of quantitation (LLOQ) is the lowest amount of analyte in a sample which can be quantified reliably, with an acceptable accuracy and precision (see Accuracy and Precision). **Proposed change:** The LLOQ corresponds to the lowest calibration point which should be adapted to expected concentrations and to the aim of the study.	The text has been revised for clarity.
151-153	48	Comment: In case of expected concentrations are not available at the time of the validation study, is there a suitable / expected LLOQ range for the toxicology studies?	Not agreed. This may be study specific.
151-153 317-365	48	Comment: The limit of detection (LOD) is not mentioned in this guideline especially for ligand-binding-assay. Indeed, for nearly every ligand-binding-assay, our sponsors request the assessment of a LOD (or « cut off » level for immunogenicity assays) below which a sample is detected as negative and above which a sample is detected as positive **Proposed change (if any):** Include the method and the acceptance criteria for the assessment of the detection limit in ligand-binding-assay	Immunogenicity assays are not covered by this guideline. In general the qualitative assays are not covered by the guideline. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
154 - 159	24	Comment: We would like to ask EMA to not define the number of replicates per concentration. We would ask EMA to remove (and IS) from the sentence since addition of IS is not part of the preparation of calibration standards. **Proposed change (if any):** We would propose to rephrase to "for each analyte, you need at least one calibration curve".	There is not such a statement here; line 158/159 does not imply a single set of samples. IS is deleted as this can be seen as the request that if bulk is prepared this also applies for IS. The proposed changed has not be taken into account, as it is considered that the text is sufficient clear.
154-181	2	Comment: • Lines 160-163 suggest knowing of the range of the method before validation. It will be very helpful for the bioanalyst to know the range before validation, but is not practical for all situations. The range can also change from study to study. Therefore, it should not be the requirement to know study concentration range for starting a validation, but that during analysis by chromatographic assays, the calibration range and/or QC concentrations should suitably cover the concentrations observed in the study. • Clarification and differentiation is needed for this section between chromatographic and ligand-binding assays. • What is the purpose of freshly prepared calibration curves especially when stability is known? Also, STDs and QCs should be treated the same way as study samples. Therefore, at least one freeze-thaw is expected. And standards and QCs should be treated the same way during validation as they will be treated during samples analysis. Hence one-freeze thaw is more appropriate than freshly prepared.	The text has been revised by adding 'Ideally'. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. First of all the guideline recommends and does not require this. In case bulk calibrators are prepared together with QC and stored, possible unexpected issues during storage will not be recovered as both calibration standards and QCs are affected. However the text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		• Define "Freshly prepared". How are the fresh calibration curves prepared and qualified to ensure that the preparation of each fresh curve is correct? Should they be prepared starting from fresh weighing of reference compounds (after comparing 2 independent weighings) or from analytical solutions and stored for some specified time. If prepared from analytical solutions, then the same question on stability exists for the analytical solution as for plasma. Also, multiple fresh preparations of calibration curves can introduce a variability associated with preparation and therefore the assay variability may be over-estimated. (Note: This topic would require a good scientific debate and therefore should not be suggested in a regulatory document. The preparation of daily calibration standards have their own set of problems and can not always be considered better than preparing bulk calibration standards when the analyte is stable). **Proposed change (if any):** • From line 158 remove "(and IS)". Addition of IS is part of performing the analytical method • Line 167: For chromatographic methods it should be clarified that "response" includes ratios with IS • Clarify "adequately": add selection of weighing factor and concentration-response fit function • It should be clearly stated what is needed for chromatographic methods and LBA, e.g., o Add: "A minimum of six calibration concentration levels should be used, excluding the blank sample (matrix	The text has been revised by adding 'freshly spiked'. IS is deleted as this can be seen as the request that if bulk is prepared this also applies for IS. It is considered obvious that the response includes ratios with IS, if applicable. The proposal is not deemed appropriate for a guideline but should be addressed in a SOP. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		without analyte for ligand binding assays and processed matrix sample without analyte and without IS for chromatographic methods) and a zero sample (processed matrix with IS, only for chromatographic methods). o Format the Guidance to contain parallel statements applicable for LBA and for small molecule. o For LBAs at least 6 calibration curves should be evaluated. o Line 174: For LBAs Accuracy of the back calculated concentrations should be within ±20%, except for LLOQ which should not exceed 25%.	
154-181	13	Comment: According to EP Chapter 2.7 three to four dilution levels are required for the reference and the samples. This is a discrepancy to the six calibration concentration levels required in the guideline and also to the five concentration levels required by the ICH guideline. The calibration as described is typically used for chromatographic methods, for the biological assays mentioned in EP Chapter 2.7 dilution series of a reference standard are performed (dilution of the matrix) and a reference curve is calculated (see general remarks).	Biologically assays are out of the scope of this guideline
154-181	21	It should be clarified whether it is acceptable to use 2 calibration curves (one before and one after the samples). In addition, it should be acceptable to use calibration curves that are not freshly prepared as long as stability over the period of use is demonstrated (For example as frozen matrix). **Proposed change (if any):** We suggest to add "It is acceptable to use 2 calibration curves (one before and	The text has been revised and indicates the possibility to use calibration standards in replicate.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		one after the samples) and to use calibration curves that are not freshly prepared as long as stability over the period of use is demonstrated".	
154-181	23	It should also be allowed to use 2 calibration curves (one before and one after the samples; it is not clear from the text if this is allowed). In addition, it should be allowed to use calibration curves that are not freshly prepared as long as stability over the period of use is demonstrated.	The text has been revised and indicates the possibility to use calibration standards in replicate.
154-181 Calibration curve	36	**Comments** Clarity is sought on the use of calibration curves not freshly prepared if stability over the use period is demonstrated. In particular if this is acceptable in partial or revalidation circumstances. Clarity is sought as to the usage of duplicate calibration standards to provide a single calibration regression.	First of all the guideline recommends and does not require this. In case bulk calibrators are prepared together with QC and stored, possible unexpected issues during storage will not be recovered as both calibration standards and QCs are affected. However the text has been revised for clarity. The text has been revised and indicates the possibility to use calibration standards in replicate.
154-181	52	Comments: According to EP Chapter 2.7 three to four dilution levels are required for the reference and the samples. This is a discrepancy to the six calibration concentration levels required in the guideline and also to the five concentration levels required by the ICH guideline. The calibration as described is typically used for chromatographic methods, for the biological assays mentioned in EP Chapter 2.7 dilution series of a reference standard are performed (dilution of the matrix) and a reference curve is calculated (see general remarks).	Biologically assays are out of the scope of this guideline.
155	23	"response of the instrument" is mentioned. however, analytes can be evaluated by counting distribution in situ in tissues. **Proposed change (if any):** Review wordings to include other possibilities of quantification.	This is not within the scope of this guideline. Furthermore, the text regarding C^{14} analysis has been deleted.
155-159	17	Comment:	The proposal is not deemed appropriate for a guideline but

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		On the one hand, the guideline defines how to prepare certain concentrations in a matrix. On the other hand, there is no definition how to generate these matrices. Is it preferable (e.g. at sera) to use a pool of batches? What is recommended when the matrix is changed?	should be addressed in a SOP. A pooled batch is acceptable. In case the matrix is changed partial validation is needed.
155, 168	11	Comment: add "concentration of" before "analyte"	The text has been revised accordingly.
156-158	45	1. Unclear if the reference to "same" matrix means disease state serum or just same matrix (e.g. human serum). 2. In addition, for Ligand Binding Assays, samples are generally diluted and therefore calibrator curves generally have a percentage of serum equivalent to the minimum required dilution vs. diluting a calibrator curve in 100% matrix. 3. Different matrices for calibration standards and study samples are allowed if validated	1. It means the same matrix, unless during validation it appeared that there is a matrix effect for instance for disease state serum. 2. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. 3. This in line with other guidance.
158	8	Comment: States that calibration curve should be spiked with IS. Typically, IS is not added during calibration curve preparation, but during the extraction procedure. **Proposed change (if any):** Remove statement	IS is deleted as this can be seen as the request that if bulk is prepared this also applies for IS.
158-159	8	Comment: In the case of multi-analyte studies, it is unclear if both analytes can be spiked in the same calibration curve.	Yes, the text indicates one calibration curve and not one set of calibration samples.
158 - 159	20	Comment: Duplicate calibration curves are widely used to improve method ruggedness and to minimize the need for sample reanalyses resulting from the loss of one or more calibration standards **Proposed change (if any):** Recommend including the possibility of using duplicate curves and revising this sentence to state that a <u>minimum</u> of one calibration curve must be included in each batch	The text has been revised and indicates the possibility to use calibration standards in replicate.
158-159	37	Comment: This point is not clear. Is one calibration curve for two analytes (two analytes in each calibration	The guideline indicates one calibration curve and not one set

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		sample) acceptable? If not there may be software problems to calibrate instruments for automatic concentration calculation. Is really only one calibration curve in method validation necessary? It should be one calibration curve for each analytical run during validation and study samples analysis. **Proposed change (if any):** In any analytical run during method validation and study samples analysis, there should be one calibration curve for each analyte studied. It is possible to use calibration samples consisting of more than one analyte.	of calibration samples.
158-159 380	47	Comment: Regarding the reference to "one calibration curve" or "a set of calibration standards." The use of a single calibration curve is acceptable for immunoassays (e.g., ligand-binding assays), but current practice in industry is to bracket an analytical run with calibration curves (therefore using two calibration curves) when conducting LC/MS methods. The importance of bracketing samples in LC/MS methods has relevance considering the applicable technology; namely, the pre-analytical extraction associated with sample preparation, reconstitution of analyte(s) and internal standard(s) (stability), chromatographic resolution, and sample ionization (differential between analyte(s) and internal standard and ion transition stability). These factors can be assumed to be acceptable in an analytical run when the acceptance criteria for both calibration curves are met. It may be that this statement is a carryover from the 2001 FDA guidance document, but there is certainly sufficient acceptance within industry for the use of bracketed calibration curves with LC/MS assays. This can be considered more relevant now given multiple LC/MS techniques in addition to triple quadrupoles in use, including hybrids, and the use of mass spectrometry being extended into macromolecular drug candidates, biomarker assays, metabolism, and drug-	The text has been revised and indicates the possibility to use calibration standards in replicate. Immunoassays are out of the scope of this guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		drug interaction studies. **Proposed:** It is appreciated that given the variations and constant advances associated with immunoassays and the pharmaceutical focus on macromolecule medicines, that flexibility is provided in the guidance to the laboratory for the use of and number of calibration curves associated with these assays. However, the guidance regarding the acceptable number of standards, as it reads, allows for the use of a single calibration curve within LC/MS methods. Since there is extensive use and justification within industry for the use of two calibration curves, which bracket an analytical run, with methods that utilize LC/MS, it is recommended the guidance make this distinction.	
160	26	For first-in-human, rising-dose studies the concentration range expected is sometimes difficult to predict from nonclinical studies. This would generally lead the analyst towards a wider calibration range than may be subsequently employed.	The text has been revised by adding 'Ideally'.
160-163	21	Although it is acknowledged that assays should be developed with a calibration range covering anticipated study sample concentrations, definition of the range prior to studies is difficult and can also change from study to study. Therefore, it should not be a requirement to know the study sample concentration range before starting a validation, but rather have a reasonable estimate. Moreover the concentration range is considered valid as validated and study samples could be and are always in case of ELISA diluted as needed to meet the calibration range.	The text has been revised by adding 'Ideally'.
160 - 163	24	Comment: We would like EMA to consider to delete the first sentence of the paragraph. **Proposed change (if any):** Please exchange the last sentence by "The range should be established to allow	The text has been revised by adding 'Ideally'. The text has been revised including the proposed text.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		adequate description of the pharmacokinetics of the analyte of interest"	
160-163	45	1. Definition of range in case of First in Human study difficult, why not using a generic range and sample dilution which is as part of the validation anyway. It will be very helpful for the bioanalyst to know the range before validation, but is not practical for all situations. 2. The range can also change from study to study. Therefore, it should not be the requirement to know study concentration range for starting a validation, but suggestion should be made on the calibration range or concentrations of the QCs during a study analysis. 3. it is not likely that in LBA the calibration range would cover the sample concentration range. e.g for mAb drugs with Cmax >100 µg/ml and sensitivities <200 ng/ml This can only be achieved by sacrificing sensitivity or validating multiple calibration range 4. In LBA, the LLOQ and ULOQ are typically not defined as the lowest and highest calibration standard, respectively. In fact, standard curve calibrators should span the anticipated LLOQ and ULOQ 5. Concentration range is valid as validated, study samples could be and are always in case of ELISA diluted as needed to meet the calibration range. 6. Calibration range selection should be guided rather by the LLOQ needed	1. The text has been revised by adding 'Ideally'. 2. This is covered in the section 5.3 Calibration range. 3-5. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. 6. This comment is unclear.
163	14	Comment: It is not clear what scientific information is. **Proposed change:** In the context of pharmacokinetics a better wording	The text has been revised according to the proposed text.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		would be: "*The range should be established to allow adequate description of the pharmacokinetics of the analyte of interest*".	
163	34	Comment: For ELISA, it is the assay itself which defines the range rather then the study.	This is agreed. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
164	30	The number of calibration levels actually depends on the analytical task, see e.g. Regression and calibration methods for analytical separation techniques Part 1: Design considerations K. Baumann, H. Wätzig Process Control and Quality 10, 59-73 (1997) Sometimes three levels suffice.	This applies to bioanalysis and not to quality issues, for which different range and number of calibration standards may be needed.
164	34	Comment: Does that mean that LLOQ and ULOQ can be included in the six calibration standards?	Not agreed, the text indicates that LLOQ and ULOQ should be included.
164-166	27	Comment: In the ICH Q2(R1), "for the establishment of linearity, a minimum of 5 concentrations is recommended", which is different from the recommendation of this guideline, lines 164-166, where "A minimum of six calibration concentration levels should be used, excluding the blank sample (processed matrix sample without analyte and without IS) and a zero sample (processed matrix with IS)" **Proposed change (if any):** Harmonize	The ICH Q2(R1) has a different scope; ICH Q2 is referring to quality issues.
164-166	45	For linear 6 calibrators are fine, for non-linear functions in LC-MS over larger concentration range more levels needed in order to define the response function better (Quadratic regression addition of calibration points?)	The guideline clearly indicates that 6 is a minimum. More calibration standards may be added.

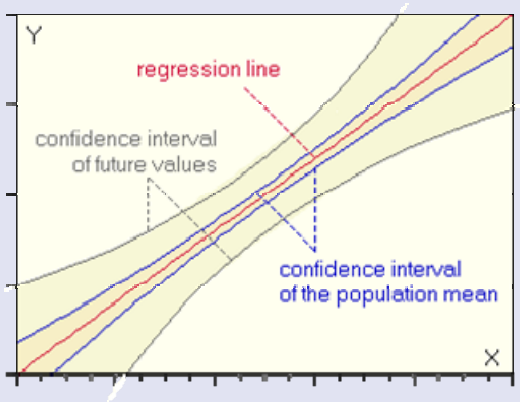
Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
164-166 168-169	47	Comment: Regarding the discussion of calibration curves, blank samples, and zero samples. Ligand-binding assays (LBAs) are not discussed here, although the attempt of the guidance appears to incorporate LBAs. **Proposed:** Recommend that this section discuss the use of anchor points and that anchor points must not be used in the calculation of the calibration curve. This should also be considered in the context of calibration curve requirements in lines 345-349.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
164 175-176 380 402	39	Is there any scientific justification of the minimum number of six calibration levels? According to our opinion the number of calibration standards depends a) on the range of calibration curve, e.g. we would require more calibration standards for 1-1000 ng/ml curve than for 1-10 ng/ml b) on the shape of calibration curve. We would require more points for parabolic calibration curve than for linear one. **Proposed change (if any):** Delete the strict requirement of at least six calibration levels.	The guideline clearly indicates that 6 is a minimum. More calibration standards may be added.
167-168	45	1. No sound statistical criteria for the definition of the best relationship concentration-response is suggested. 2. Blank matrix subtraction is necessary for log-log and calibration curve parameters are affected	1. Which function can be used is up to the analist. The proposal is not deemed appropriate for a guideline but should be addressed in a SOP. 2. This comment is unclear.
167-169	17	Comment: A relationship which can "simply" and "adequately" describe... is proposed. It would be helpful to specify what is meant by "simply" (e.g. linear or square fit)	We state 'simply' to prevent that too complex formulas are used, without a scientific justification.
167 - 169	20	Comment: How will the phrase "…simply and adequately…" be defined? Some labs use a quadratic fit	We state 'simply' to prevent that too complex formulas are used, without a scientific justification.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		as a default for all methods; would this approach meet the simple and adequate criteria? Generally the choice between linear unweighted, linear 1/x, linear $1/x^2$, quadratic, power fits, etc. are made subjectively, based on analyst experience and preference On several occasions we have been asked by reviewers to develop objective criteria for determining the simplest acceptable regression model for a bioanalytical method. The intent was that different scientists reviewing the same data set would choose the same regression model **Proposed change (if any):** If the intent of the EMA is to require justification that the chosen algorithm is the most simple and adequate model, then a recommendation should be included stating that a decision tree with objective criteria should be implemented for choice of regression algorithm. If this is not the intent, then no changes are necessary	No justification is requested.
170	47	Comment: Regarding the statement "The calibration curve parameters should be submitted." **Proposed:** Recommend that "reported" be used rather than "submitted."	The text has been revised accordingly.
170-173	8	Comment: The curves that we are required to present are unclear. Do we need to present any 3 curves? The curves used for the precision and accuracy evaluation? All the available (or acceptable) curves in the validation?	The text has been revised for clarity, stating that all the available curves in the validation, with a minimum of 3, should be reported.
170 - 173	24	Comment: We would like EMA to change the wording from submitted to reported. **Proposed change (if any):** This section generally relates to the reporting section and please move the paragraph to the reporting section.	The text has been revised accordingly.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
170-172	37	Comment: Precision of measurement results of calibration standards seems to be necessary. **Proposed change (if any):** In addition, the back calculated concentrations of the calibration standards should be presented together with the calculated mean accuracy and precision values (see ~~definition of Accuracy~~ below).	For calibration standards, precision is not required.
170-173	45	1. Why should three calibration curves be presented,? 2. Please clarify that the following statement applies to small molecule (chromatography methods) only "At least 3 calibration curves should be evaluated"	1. Reporting of the results of at least 3 calibration curves is a regulatory set minimum criteria. 2. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
172	8	Comment: It is necessary to present the calculated mean accuracy values, however, there is no criteria stipulated.	As indicated in line 174 – 178, the calibrators should be excluded with >15% inaccuracy; meaning that the remaining calibrators will fit within 15%.
172-173	19	What do you mean on "At least 3 calibration curves should be evaluated"? Should these calibration curves be derived from 3 different validation batches (using 3 linear fittings) or should 3 sets of calibration standards be analysed together on one occasion (using one linear fitting)? This is not evident. **Proposed change (if any):** Please clarify.	The text has been revised for clarity, stating that all the available curves in the validation, with a minimum of 3, should be reported.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
172-173	22	Comment: Point 4.1.4 Calibration curve It is unclear what is meant by "At least 3 calibration curves should be evaluated". Should these calibration curves be derived from 3 different validation batches (using 3 linear fittings) or should 3 sets of calibration standards be analysed together on one occasion (using one linear fitting)?	The text has been revised for clarity, stating that all the available curves in the validation, with a minimum of 3, should be reported.
172-173	42	Comment: inter day results could not be obtained **Proposed change (if any):** At least 3 calibration curves should be evaluated in a span of 2-3 days , not in a single day. (To have a inter day concentrations data since the subject sample analysis could not be done in a single day.)	The text has been revised for clarity, stating that all the available curves in the validation, with a minimum of 3, should be reported. Furthermore accuracy and precision of QCs should be assessed over more than 1 day, meaning that calibration curve results over more than 1 day will be available.
172-173	47	Comment: Regarding the statement "At least 3 calibration curves should be evaluated." **Proposed:** Recommend clarifying the necessity of only three calibration curves being evaluated. Specifically, all curves should be evaluated, however at least three calibration curves should be reported (for traditional LC/MS or conventional chromatographic analyses). Validations of other types of methodologies, such as ligand-binding assays, often include calibration curves of n=6, so further distinction is recommended.	The text has been revised for clarity, stating that all the available curves in the validation, with a minimum of 3, should be reported. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
172-173 175-178	23	"At least 3 calibration curves should be evaluated." "In case a calibration standard does not comply with…" During the analysis of study samples it is accepted the truncation of calibration curve due to rejection of LLOQ or ULOQ calibration standard. If this idea is maintained, this possibility should be reflected also during the validation.	The text has been revised indicating the criteria to be applied. Also with regard to the LLOQ and the ULOQ.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): At least 3 calibration curves with LLOQ and ULOQ should be evaluated	
172 -173 205-208	48	Comment: The assessment of the between and within accuracy and precision for the validation study with only 3 runs seems not to be enough especially for ligand-binding-assay with the 6 * 6 format seems the more appropriate. For LC/MS method, the 5*5 format is currently used. The 3*5 should be adopted if the EMA maintain this format. **Proposed change (if any):** Suggestion of two different format depending on the nature of the study	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
174-	51	Comments: Often, two curves may be used, you describe accuracy but not precision. Add requiment for precision if more than one curve is used. **Proposed change (if any):** Add statement describing precision for point if one or more curves are used	For calibration standards, precision is not required.
174	51	Comments: Add the actual formula for %Bias or %RE and %CV at the first description. It sound elementary but some folks in the ligand binding arena are using SD of mean/nominal as CV rather than SD of Mean/Mean. In actual fact they are using both calculations! **Proposed change (if any):** Add description of the formulae at first description.	Calculation of accuracy and precision is a basic (simple) calculation and it is considered that adding formulas are not needed.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
174-175	37	Comment: It is scientifically justified that two points of calibration curve are determined with higher uncertainty (see for example graphical confidence interval for calibration curve): LLOQ and ULOQ.  It should be proposed to widen acceptance criteria on both ends of calibration curve (acceptance criteria for ULOQ should be the same as for LLOQ). **Proposed change (if any):** The back calculated concentrations of the calibration standards should be within ±15% of the nominal value, except for the LLOQ and ULOQ for which it should be within ±20%.	The basic requirements are still 15% at ULOQ and 20% at LLOQ.
174-178	5	Comment: In chapter 5.2 (Acceptance criteria of an analytical run) it is mentioned that the criteria for decision to exclude calibration standards or not should be pre-defined in a SOP, and should be independent from the results of the QC samples. This procedure should also be included in this chapter. There should be no difference in demand on the calibration curve of validation and on the calibration curve of the analytical run.	The text is considered sufficiently clear.
174-178	19	There are no criteria for the calibration standards per concentration level. With the present wording it may occur that all of the replicates from the same	In case replicates are used, this is agreed. The text has been revised accordingly.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		concentration level are rejected. **Proposed change (if any):** The criteria (±15% and ±20%) should also be fulfilled for at least 50% of the calibration standards tested per concentration level.	
174-178	21	Distinguish here between LC-MS and ELISA acceptance criteria.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
174-178	22	Comment: Point 4.1.4 Calibration curve There are currently no criteria for the calibration standards per concentration level. With the present wording we see a risk that all of the replicates from the same concentration level could be rejected. **Proposed change (if any):** We propose that the criteria (±15% and ±20%) should also be fulfilled for at least 50% of the calibration standards tested per concentration level.	In case replicates are used, this is agreed. The text has been revised accordingly.
174-178	45	Distinguish here between LC-MS and ELISA acceptance criteria	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
174-178	49	There is no criteria for r or r^2. **Proposed change (if any):** Acceptance criteria for r or r^2 should be included.	This is considered not relevant.
175	23	"At least 75% of the calibration curves standards with a minimum of six, must fulfil this criterion". Does that mean that the curve is composed of 6 points and 4 are acceptable or do we need 6 acceptable points so at least	The guideline clearly indicates that 6 is a minimum for acceptable calibration standards. More calibration standards may be added.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		8 in the curve?	
175-176	42	Comment: **Proposed change (if any):**minimum of six including the lowest and highest point (These are anchoring points and hence should not be sacrificed.)	The text indicates that LLOQ and ULOQ should be included. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
176	23	At least four out of six non-zero standards should meet the above criteria, including the LLOQ and the calibration standard at the highest concentration. Excluding the standards should not change the model used (FDA 2001).	The guideline clearly indicates that 6 is a minimum for acceptable calibration standards. More calibration standards may be added. **Four out of six is a criteria for QC samples.**
176	23	It would be useful to ensure coherence with the FDA requirements.	The EMA considers harmonisation an important issue. Harmonisation issues have been taken into account as much as possible.
176 - 178	20	Comment: The issue of a failed LLOQ or ULOQ standard has not been addressed. If a singlet curve is used, should the curve be truncated and retained in the validation or should this batch be rejected from the validation? **Proposed change (if any):** Recommend that if a failed LLOQ or ULOQ occurs in a singlet curve validation that the batch should be rejected from the validation, the source of the failure be determined, and the method revised (if necessary). If the next validation batch also fails, then the method must be revised before restarting validation	The text has been revised including the proposal.
176-178	47	Comment: Regarding the rejection of calibration standards. **Proposed:** Recommend that this section be re-written to incorporate ligand-binding assays, as well as the importance of documenting the acceptability of standard	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. Truncation is not requested to be validated; it is agreed that

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		curve truncation within the validation report (and method). Additionally, the allowance (as an a priori procedure) for calibration curve truncation and acceptable practices should be included here.	during study sample analysis truncation may be necessary, however it is considered that the curve is still applicable as long as they fulfil the accuracy criteria as requested.
176 and 402	4	Comment: Line 164 states, "A minimum of six calibration levels". This needs clarification since many organisations use duplicate calibration standards. **Proposed change (if any):** "At least 75% of the calibration standards with a minimum of six calibration levels, must fulfil this criterion."	The guideline clearly indicates that 6 is a minimum for acceptable calibration standards. More calibration standards may be added. The text has been revised for clarity.
176, 402, 425	11	Comment: add "concentration levels" after "six"	The text has been revised for clarity.
Insertion after the lines 178	42	Comment: For the combination of drugs and the metabolites, the procedure for the construction of CC, concentration data is not given **Proposed change (if any):** For the combination of drugs and metabolites, the calibration points also should have mimic in this fashion to simulate the study samples (This is very much essential,since we are now having the combination of drugs in bioequivalce studies.)	It is considered that this is sufficient covered
179-181 and further in the text	1	The proposal to use freshly (daily) prepared calibration curves during the validation needs a clarification on what is intended: new weighing, dilutions from available stocks or what else. It would be also important to clarify if such approach applies also to the analyses of a study. In such case it is noteworthy that an increase of the results variability may derive by curve preparation procedures; this fact may influence the study results.	The text has been revised to clarify this issue. It indicates now 'freshly spiked'.
179-181	6	Comment: Once it is clear from stability data that the analyte is sufficiently stable in the matrix of interest the freshly	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		prepared calibration curves during the validation are not necessary, except for long lasting tests – e.g. long-term stability. **Proposed change (if any):** Either delete this paragraph or specify tests for which freshly prepared standards are strictly required.	
179-181	7	Comment: It should be allowed to use calibration curves that are not freshly prepared as long as stability over the period of use is demonstrated.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179-181	8	Comment: If long-term stability data demonstrates that the calibration curves are stable for the duration of the validation, why can't bulk preparations be used? It reduces the amount of time and resources that what is required for freshly spiked calibration curves. **Proposed change (if any):** Remove this requirement	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179-181	14	Comment: Preparation of fresh calibration standards can be limited to the first validation of a method. If reliable stability data on the analyte in the respective matrices are available during development of a drug, further validations can also be performed with frozen calibration samples.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179-181	18	Comment: Once it is clear from stability data that the analyte is sufficiently stable in the matrix of interest the freshly prepared calibration curves during the validation are not necessary, except for long lasting tests – e.g. long-term stability. **Proposed change (if any):** Either delete this paragraph or specify tests for which freshly prepared standards are strictly required.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179-181	21	It should be allowed to use calibration curves that are	The text has been revised, including the use of calibration

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		not freshly prepared as long as stability over the period of use is demonstrated. **Proposed change (if any):** Although ~~it may be clear from stability data that the analyte is sufficiently stable in the matrix of interest,~~ it is recommended ~~that~~ to use freshly prepared calibration curves ~~are used~~ during validation of the bioanalytical method*, it is allowed to use calibration curves that are not freshly prepared as long as stability over the period of use is demonstrated.*	curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179-181	23	Why should we use freshly prepared calibration curve if stability is demonstrated in a frozen matrix? If stability in a frozen matrix has been established, the validation could then be conducted using frozen calibration curves so this section should be modified in that sense. What does "freshly prepared calibration curve mean"? Is it necessary to prepare the calibration samples freshly each day? If calibration samples are prepared within one week before the validation study and then stored in single use aliquots: Is it necessary to compare these samples once with a freshly (not stored) set of calibration samples or QC samples even if stability has previously been demonstrated?	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179 - 181	24	Comment: We believe and experienced that stored samples can be used for validation as long as stability has been demonstrated. **Proposed change (if any):** We suggest the addition of the following: However, for analytes for which stability has been demonstrated, calibration standards can be stored and used over the period and conditions of demonstrated stability.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179-181	37	Comment: In case of no stability issues there is no reason to use freshly prepare calibration standards. No scientific justification, no improvement in measurement results quality. It seems to be an overregulation.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): ~~Although it may be clear from stability data that the analyte is sufficiently stable in the matrix of interest, it is recommended that freshly prepared calibration curves are used during validation of the bioanalytical method.~~	
179-181	39	We do not understand this paragraph. If the analyte is stable in the matrix of interest e.g. for one month, we can use up to one month old calibration samples. What could be wrong on this approach? **Proposed change (if any):** Delete this paragraph.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179-181	40	Comment: The guideline recommends using freshly prepared calibration curves during validation. I disagree with this for the following reasons: 1. The purpose of the validation is to assess the validity of the method. Therefore, validation with the same spiking is more representative. If the spiking was done fresh every day, a difference in spiking could create a false outcome, either positive or negative, therefore the inter–assay would no longer be representative of the method. 2. For stability assessment, it is more beneficial and more representative with the sample from the same spiking as well. In this case, the only variable is the stability. If fresh spiking was used, there are two variables: spiking and stability. Therefore, assessment for the fresh spiking will not be as accurate as with the same spiking. 3. To do a fresh spiking every day adds extra work without real benefit. As long as the stability of the plasma sample can be established, it would be sufficient. Of course, samples that are stored too long should not be used for validation.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Generally speaking, a method can be validated within two weeks. Therefore, I believe this is a good timeline for the sample to be considered relatively "fresh" enough. **Proposed change (if any):** Although it may be clear from the stability data that the analyte is sufficiently stable in the matrix of interest, it is recommended that the calibration curves used during the validation of the bio-analytical method should be prepared no longer than two weeks.	
179-181	45	1. Lines 179-181 suggest preparation of fresh calibration curves during validation even when the analyte is stable in the matrix. This topic would require a good scientific debate and therefore should not be suggested in regulatory document. The preparation of daily calibration standards have their own set of problems and cannot always be considered better than preparing bulk calibration standards when the analyte is stable. 2. What is the purpose of freshly prepared calibration curves especially when stability is known? Conceptually, we are expecting STDs and QCs treated the same way as study samples. Therefore, at least one freeze-thaw is expected	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179-181	48	Comment: What does the draft document suggest by "freshly prepared calibration curve". Does it mean that the standard calibration (S1 to Sn) freshly prepared for each series from either: - a stock solution prepared previously (stability data should so covered the period between its preparation and its use before the end of the study) - or x (1 to n) intermediate solution (in solvent not in matrix)prepared previously (stability data	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		should so covered the period between its preparation and its use before the end of the study) are both compliant with the term "freshly prepared"? If not , this point would like to say that the weighing or the required volume sampling should be performed at each new series of the validation study. **Proposed change (if any):** Give more details about the meaning of "'freshly " prepared.	
179-181	49	This is only possible if the stock solution prepared is stable. This approach is not possible when there is a limited quantity of analyte available to prepare solutions.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179-181	50	Although it may be clear from stability data that the analyte is sufficiently stable in the matrix of interest, it is recommended that freshly prepared calibration curves are used during validation of the bioanalytical method. Whether study samples should be analyzed using freshly prepared or bulk spiking calibration curve or not, this is to be made clear to avoid any discrepancy in practical use? Details on use of freshly prepared or bulk spiking calibration standards should be incorporated in this section to avoid any ambiguity.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
179-181	51	Comments: I would add this to the description of the QC preparation during accuracy and precision runs. This exercise would demonstrate that QCs can be prepared and acceptable values returned over several days. Otherwise one single batch of QC is prepared, stored without valid stability data, and use to assess the method. **Proposed change (if any):** Add statement about freshly prepared QCs to 4.1.5., and 4.1.6	This is not required for QCs.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
180-181	29	Comment: The use of fresh calibration curves during method validation need not be recommended during method validation. Alternatively bulk spiked calibration standards can be used, during method validation (for example for establishing precision and accuracy of quality control samples), except when they are required to perform stability experiments. This can avoid errors in the preparation of solutions and spiking errors, if any. **Proposed change (if any):** The use of freshly prepared calibration curve is recommended, although not mandatory, during method validation.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
180-181	41	Comment: Use of fresh calibration curves in entire validation is contradicting the FDA Guidance, in which this is requirement in any stability experiment. **Proposed change (if any):** Ensure harmonization between guidelines.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
180-181	47	Comment: Regarding the statement that "…it is recommended that freshly prepared calibration curves are used during validation of the bioanalytical method." Current industry practice is that many laboratories will prove matrix stability early in the validation process so as to begin to provide initial data for the study samples themselves. As a result of this early stability work, these laboratories will use this data to support the use of frozen calibration standards for subsequent method validation analyses. **Proposed:** Recommend that the use of freshly prepared calibration curves in the validation not be listed as a requirement for all validation studies, but rather only for stability studies and those studies performed where frozen matrix stability has not been established.	The text has been revised, including the use of calibration curves that are not freshly spiked and prepared as long as stability over the period of use is demonstrated.
182	23	In the section 4.1.5the terms “true concentration” or “nominal value” are used. It would be useful to harmonize these terms for the whole document and to	The term ‘nominal value’ will be used.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		put in a definition for "nominal value in the definitions section (L545).	
182-198	2	Comment: • It is stated that the accuracy should be reported as percent of the nominal value whereas criteria for within/between run accuracy are given as % bias (deviation from the nominal value). • For large molecules ULOQ QC should be included for precision and accuracy. • Whilst a mid QC at around 50% provides demonstration of method P&A performance. It is not necessarily the best concentration for a Mid QC to provide a demonstration of control for the majority of real samples. **Proposed change (if any):** • For large molecules, ULOQ QC should be included for precision and accuracy. • Add ranges that better reflect expected samples - perhaps Mid QC of 30 to 50% of range OR geometric mean of Low and High QC. High QC could be 60 to 85% of the ULOQ.	Indeed, however this is considered not problematic and obvious. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. A median QC concentration around 50% of the calibration curve range is recommended. A high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations.
182-198	15	A more flexible range for the mid QC (not fixed to 50 %) should be applicable. Otherwise there would be a quite big gap between the low and mid QC.	A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations.
182-354	27	Comment: Accuracy, precision etc –Lines 182-354 **Proposed change (if any):** This should be harmonized with ICH Q2(R1) ; A plan R&R (repetability&reproducibility) integrating all the variables cited would be of interest	The ICH Q2 guideline deals with quality issue. This is a different scope.
183-198	22	Comment:	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Point 4.1.5 Accuracy Between run accuracy (QC's outliers) Bioanalysis is an experimental work and it can sometimes prove challenging to explain an anomalous result for a QC test. The EGA would support the inclusion in the guideline text of the possibility of excluding an anomalous QC value, with the provision that a reasonable explanation or justification exists for it (lab errors).	The text has been revised taken into account this issue.
183-184	37	Comment: In our opinion the definition should be consistent with definition widely used all over the world (e.g. FDA). Acceptance criteria 85-115% and 80-120% for LLOQ. **Proposed change (if any):** The accuracy of an analytical method describes the closeness of the determined value obtained by the method to the true concentration of the analyte. It is calculated as measured concentration / nominal concentration ratio (expressed in percentage).	This is considered a minor point as the requested recommendations are clear.
187-190	21	For LBA FDA guideline recommends five concentration levels: LLOQ, 3 x LLOQ, 50 % of range, 75-80 % of ULOQ and at ULOQ. Distinguish here between LC-MS and LBA	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
187-190	21	Replace suggested concentration for the medium QC to geometric mean of the low and high QC, as also indicated in FDA guidance **Proposed change (if any):** During method validation accuracy should be determined by replicate analysis using a minimum of 5 determinations at a minimum of 4 concentration levels which are covering the calibration curve range: the LLOQ, within three times the LLOQ (low QC), around ~~50%~~ *the geometric mean* of the calibration curve range (medium QC), and at about 75%	A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		of the upper calibration curve range (high QC).	
187-190	23	The levels of QCs during validation are defined 3 x LLOQ, 50% and 75% calibration range. Calibration curves with LC/MS methods are often populated asymmetrically, i.e. more calibration samples in the low range, less in the high concentration range. So the mid QC is often not placed at 50% calibration range, but around the "middle" calibration samples, i.e. if you use 8 calibration samples the mid QC is often between the 4th or 5th standard. When one rejects calibration points other than for analytical reasons, must these values be included in the cumulated statistical analysis of the standard calibration curve? If yes, how should one manage the out of specification results which could have brought about a variation in the statistical values thus calculated?	A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations. The should be reported, but not taken into consideration
187-190	45	1. No accuracy at ULOQ is mentioned here. However, samples at high concentrations are contributing to a much larger extent to the AUC rather than samples close to the LLOQ. 2. Replace suggested concentration for the medium QC to geometric mean of the low and high QC 3. Please clarify that these lines apply to small molecule (chromatography) methods only	1. The accuracy of the ULOQ should be within 15% of the nominal value. 2. A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations. 3. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
187-190; 200-208; 213-223	1	The introduction in the validation procedures of the within-run and between-run accuracy and precision at LLOQ, analyzed as replicate QC, is statistically and analytically questionable forcing the estimation of concentrations outside the calibration range (values < LLOQ).	It is considered that an independent analysis and evaluation is more adequate.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change: It would be better to carry-out this test considering the replicate LLOQ all as Calibrator 1 and not as QC.	
187-188	8	Comment: When the number of replicates required for an evaluation is given, is this the minimum number of evaluable replicates? What is one replicate is lost due to an analytical issue?	The guideline recommends 5 replicates, but does not indicate acceptability. There is no minimum of passing results.
187-190	6	Comment: We do not agree with the suggested concentration of medium QC particularly in relation with requirement to narrow calibration range in chapter 5.3. For example, with the calibration curve 1-1000 ng/ml, the concentration of low QC would be 3 ng/ml, that of medium QC 500 ng/ml and that of high QC 750 ng/ml. Thus, the concentration of medium QC would be 167 times higher than that of low QC, but high QC will be only 1.5 times higher than medium QC. According to our opinion, geometric mean of low QC and high QC is more suitable	A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations.
187-190	47	Comment: Regarding the reference to the setting of the level of the middle QC sample. There are alternative approaches to setting the level of the middle QC sample (e.g., geometric mean of the low and high QC samples) that are equally valid to the method recommended in the guidance. **Proposed:** Recommend that the guidance allows for flexibility in the approach for the determination of the placement of the middle QC sample level.	A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations.
189-190	7	Comment: Geometric mean preferred for medium QC not 50% of the calibration curve range	A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
187 - 190	10	Comment: According to the guidance accuracy (within-run and between-run) will be determined with the following samples: LLOQ, low QC, medium QC and high QC. These samples covering only 75% but not the whole calibration range. **Proposed change (if any):** We would propose to add a sample corresponding to the upper limit of quantitation (ULOQ) to cover the calibration range.	The accuracy of the ULOQ should be within 15% of the nominal value.
187 - 190	24	Comment: We would like EMA to consider the need to assess accuracy as well at the ULOQ in case of a non-linear concentration-response relationship of the assay. Also, please add that the medium QC may also be performed at the geometric mean of the LLOQ and ULOQ. The QC samples and the calibration samples should be independently prepared.	The accuracy of the ULOQ should be within 15% of the nominal value. The text has been revised and recommends that QCs and calibration standards should be prepared independently.
187-190	37	Comment: FDA requires accuracy /precision for QC samples only and so is the current practice. It was shown through years that this approach gives acceptable results. Therefore, in our opinion, EMA regulations should be similar to FDA. As stated above (see comments to lines 174-175) LLOQ and ULOQ are both determined with higher point of uncertainty. So if it is necessary to examine accuracy/precision for LLOQ so should applay to ULOQ. There is no evidence that additional experiment for precision/accuracy (LLOQ or LLOQ + ULOQ) will significantly improve measurement results quality. **Proposed change (if any):** During method validation accuracy should be determined by replicate analysis using a minimum of 5 determinations at a minimum of 3 concentration levels ~~which are covering the calibration curve range~~: within three times the LLOQ (low QC), about ~~around~~ 50% of the calibration ~~curve~~ range (medium QC), and at about 75% of the upper calibration ~~curve~~ range (high QC).	The definition of LLOQ indicated an acceptable precision and accuracy, meaning that this should be demonstrated. A alleged deviation regarding LLOQ cannot be confirmed.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
187-190	39	We do not agree with the suggested concentration of medium QC. For example, with the calibration curve 1-1000 ng/ml, the concentration of low QC would be 3 ng/ml, that of medium QC 500 ng/ml and that of high QC 750 ng/ml. Thus, the concentration of medium QC would be 167 times higher than that of low QC, but high QC will be only 1.5 times higher than medium QC. According to our opinion, geometric mean of low QC and high QC is more suitable	A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations.
188, 200, 205	40	Comment: The guideline recommends using 5 QC and LOQ samples for within and between run accuracy. I would recommend changing it to 6, as batch acceptance criteria is generally defined as 67% for QC samples with 15%. It is easier from the calculation and batch acceptance stand point. **Proposed change (if any):** Change "five samples per concentration level" to "six samples per concentration level".	This section deals with validation. Batch acceptance criteria as 67% for QC samples with 15% is applicable to an analytical run. The guideline recommends 5 as a minimum; more samples may be included.
189	9	Medium QC concentration is suggested to be "*...around 50% of the calibration curve range (medium QC).... "*. This comment needs to be qualified for ligand binding assays. **Proposed change (if any):** Suggest that, for ligand binding assays, the MQC should be selected at a concentration approximately equal to the midpoint of the dynamic range of the assay on a log scale.	A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
189-190	19	50% of the calibration curve range for medium QC is too high when a wide range is used. High QC should have higher concentration than indicated to cover the calibration range more precisely. **Proposed change (if any):** We suggest to change the	A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		text as follows: "… the LLOQ, within three times the LLOQ (low QC), around 10-50% of the calibration curve range (medium QC), and at 70-90% of the upper calibration curve range (high QC)."	
189-190	41	Comment: medium QC level should be chosen at about 50% of the range (e.g. range of 1-1000 -> QC medium at 500). This make an uneven distribution of QC's over the range. **Proposed change (if any):** It is better to use the geometric mean of the range (e.g. range of 1-1000 -> QC medium at 50-100)	A low QC at three times the LLOQ, a median QC concentration around 50% of the calibration curve range and a high QC concentration at at least 75% of the calibration curve range is recommended. The applicant is free to choose additional concentrations.
190	6	Comment: conflict with FDA high QC definition **Proposed change (if any):** at about 75-90% of the upper calibration curve range (high QC).	The alleged conflict cannot be confirmed. The text has been revised and includes now a 75% cut off value.
190	18	Comment: conflict with FDA high QC definition **Proposed change (if any):** at about 75-90% of the upper calibration curve range (high QC).	The alleged conflict cannot be confirmed. The text has been revised and includes now a 75% cut off value.
190	45	**Proposed change (if any):** Suggest that it indicate "75% of the calibration curve range" instead of "75% of the upper calibration curve range"	The text has been revised and includes now a 75% cut off value.
After the lines 190	42	Comment: We have to foresee a condition, in which LLOQ/ or ULOQ fails **Proposed change (if any):** the concentrations are chosen in such a fashion, that there should be always 2 CC points above HQC and 2 CC points below LQC. (This is very much crucial during the subject sample	This is considered not an issue to be covered in this guideline. This is left to the CRO/analists.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		analysis .When LLOQ or ULOQ fails, we can truncate the CC and since all the 3 QCs shall be again within the modified range. Otherwise, one QC shall be out of range resulting, only 2 QC in the range, which is not fair.)	
After the lines 190	42	Comment: The stability of the samples in case of combination of drugs is not specified **Proposed change (if any):**The bench top stability and refrigeration stability, freeze thaw stability and Long term stability of the samples in the matrix should include the both the drugs and/ or active metabolite to have a exact simulation of study samples (This is to simulate the study sample scenario)	Section 4.1.9 Stability, has been revised and recommends: 'In case of a multi-analyte study and specific for bioequivalence studies, attention should be paid to stability of the analytes in the matrix containing all the analytes.'
191 -195	24	Comment: **Proposed change (if any):** The last sentence should be extended to "during routine analysis".	For clarity line 194-195 has been deleted.
191-195	45	Please remove the label "QC" from the text and replace by "validation samples" (they could be to a great extent identical but not necessarily).	It is considered that the use of 'QC' instead of 'validation samples' does not lead to confusion. Therefore the term "QC' has been kept.
191-198	21	Please remove the label "QC" from the text and replace by "validation samples" (they could be to a great extent identical but not necessarily). **Proposed change (if any):** The ~~QC~~ *validation* samples are analysed against the calibration curve, and the obtained concentrations are compared with the nominal value. The accuracy should be reported as percent of the nominal value. Accuracy should be evaluated for the values of the ~~QC~~ *validation* samples obtained within a single run (the within run accuracy) and in different runs (the between-run accuracy). The latter will support the accuracy over time. To enable evaluation of any trends over time within one	It is considered that the use of 'QC' instead of 'validation samples' does not lead to confusion. Therefore the term "QC' has been kept.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		run, it is recommended to demonstrate accuracy of ~~QC~~ *validation* samples over at least one of the runs with a size equivalent to a prospective analytical run.	
193-194	47	Comment: Regarding the evaluation of inter-run accuracy using QC samples "…in different runs." **Proposed:** Recommend that this be clarified since accuracy data is usually reported for all runs within a method validation rather than simply "different runs."	It is not understood why the data of all runs should not be reported.
194-195	37	Comment: This sentence is unnecessary. **Proposed change (if any):** ~~The latter will support the accuracy over time.~~	For clarity line 194-195 has been deleted.
196	23	It is recommended to demonstrate accuracy of QC samples over at least one of the runs with a size equivalent to a prospective analytical run. What is the usefulness? The bench-top stability and the auto-sampler stability can cover this point.	This has notting to do with stability, but with accuracy and precision during a complete run. It ensures that the method is also accurate and precise in case a full run is applied. Still it is recommended and not a must.
196-198	6	Comment: This would represent a substantial burden to the laboratory. It will increase costs and time required for validation without any statistically evaluable information. Let us suppose that one analytical run contains 200 samples including 20 QC samples. The last two high QC samples will be out of the allowed limit, the other samples will be within the limits. How to evaluate this information? Is it random fluctuation or does it mean that the accuracy is compromised at high concentrations in last samples in the run? It is not possible to obtain the answer from a single run. **Proposed change (if any):** Delete this paragraph.	The guideline recommends it and does not require it. Furthermore, it is not ethical to repeat a study in case no study sample is left for reanalysis, because validation was not done properly.
196-198	7	Comment:	The guideline recommends it and does not require it.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		It is a pure business risk if one discovers in the first study batch that the batch is not valid due to a negative trend. **Proposed change (if any):** Please delete the entire sentence	Furthermore, it is not ethical to repeat a study in case no study sample is left for reanalysis, because validation was not done properly.
196-198	14	Comment: We recommend deletion of the paragraph, since the resulting batch sizes from clinical studies from clinical development drugs are not foreseeable during method validation. Thus, the robustness of a method can only be assessed from real studies and not with an artificial run. **Proposed change:** Please delete the paragraph.	The guideline recommends it and does not require it. Furthermore, it is not ethical to repeat a study in case no study sample is left for reanalysis, because validation was not done properly.
196-198	15	This should not be a requirement since the robustness of a method can depend on different parameters which may change during the conduct of the study, e.g. column. It is not a guarantee that a longer run with study samples will not fail depending on the life time of e.g. the column. **Proposed change:** Please delete the sentence.	The guideline recommends it and does not require it. Furthermore, it is not ethical to repeat a study in case no study sample is left for reanalysis, because validation was not done properly.
196-198	21	Please delete entire sentence. It is a pure business risk if one discovers in the first study batch that the batch is not valid due to a negative trend.	The guideline recommends it and does not require it. Furthermore, it is not ethical to repeat a study in case no study sample is left for reanalysis, because validation was not done properly.
196 - 198	24	Comment: **Proposed change (if any):** We would like to ask EMA to change the sentence to: "To enable evaluation of any trends over time within one run, accuracy of QC samples over at least one of the runs with a size equivalent to a prospective analytical run should be considered"	The text has been partially revised. Furthermore, precision should also be taken into account. This has been included now.
196-198	37	Comment: Generally, as described in lines 468-471, the	The guideline recommends it and does not require it.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		validation samples do not mimic the actual study samples. If there are trends in time, they will be observed during analysis of study samples. So there is no scientific justification for described experiment, it is overregulation. **Proposed change (if any):** ~~To enable evaluation of any trends over time within one run, it is recommended to demonstrate accuracy of QC samples over at least one of the runs with a size equivalent to a prospective analytical run.~~	Furthermore, it is not ethical to repeat a study in case no study sample is left for reanalysis, because validation was not done properly.
196-198	39	This would represent a substantial burden to the laboratory. It will increase costs and time required for validation without any statistically evaluable information. Let us suppose that one analytical run contains 200 samples including 20 QC samples. The last two high QC samples will be out of the allowed limit, the other samples will be within the limits. How to evaluate this information? Is it random fluctuation or does it mean that the accuracy is compromised at high concentrations in last samples in the run? It is not possible to obtain the answer from a single run. **Proposed change (if any):** Delete this paragraph.	The guideline recommends it and does not require it. Furthermore, it is not ethical to repeat a study in case no study sample is left for reanalysis, because validation was not done properly.
196-198	45	1. Robustness testing not needed because in study analysis study samples are always bracketed by QCs, and only applicable to LC-MS 2. Please clarify that these lines apply to small molecule (chromatography) methods only	1. The guideline recommends it and does not require it. Furthermore, it is not ethical to repeat a study in case no study sample is left for reanalysis, because validation was not done properly. 2. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
196-198	47	Comment: Regarding the "evaluation of trends" using accuracy of	The guideline recommends it and does not require it. Furthermore, it is not ethical to repeat a study in case no

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		QC samples within one run. Although we agrees that the evaluation of trends is critical for evaluating the robustness of a method, and the acceptance of method validation runs should not be evaluated with a 'check box' approach, assessing trends over a single run is inadequate. For instance, in the case of carryover only just meeting acceptance criteria, or when the QC samples at the beginning of an analysis quantitate higher than nominal whereas the QCs at the end of an analysis read lower than nominal, single instances do not shed light on potential carryover or the differences between poorly prepared QC samples or inherent variations of matrix components that potentially affect the method (the likelihood of which will impact incurred samples). **Proposed:** Recommend assessment of the impact of the length of a run, particularly in the cases of LC/MS methods when a run constitutes more than a single sample block, or when there is a loss of signal response over the course of a run that is independent of analog internal standard stability, or in the case of LBA assays, fast enzymatic activity. This should be clarified in this paragraph of the guidance. Validating the length of a run, (thereby assessing the impact of sample size), is prudent. As this paragraph reads, other trends are also suggested to be analyzed in a single run. This may not be a good practice since the trend over several runs would be more conclusive. This latter practice is worth noting since the guidance and/or White Papers can often be used as the sole instructions for validating a method that may lead to complacency regarding the evaluation of the data, even in runs that pass.	study sample is left for reanalysis, because validation was not done properly. More runs may be applied; 1 run is set as a recommendation by this guideline. Additional tests may be applied.
199-203	2	Within-run accuracy Comment: • Text does not clarify that acceptance criteria are for small molecule only.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		• Also, within-run accuracy (N=1) is addressed by between-run accuracy and between-run portion of the "between-run" precision. Separate requirement for the "within-run accuracy" might not add more information.	Separation of within run and between run accuracy and precision is considered more clearer.
199-203	5	Comment: For accuracy acceptance criteria are defined for the mean accuracy. Accuracy investigation should be performed with at least five samples per concentration. During preparation of the samples a sample preparation error can occur, during measurement a measurement error can occur or individual samples can deviate e.g. more than 15 %. Therefore, beside the mean accuracy criteria additional criteria for the individual results of the QC samples should be defined e.g. the acceptance criteria of an analytical run (67 %; see chapter 5.2, line 412-413).	No additional criteria have been included. Furthermore the guideline indicates that determination of accuracy and precision should include all results obtained except those cases where errors are obvious and documented.
199-203	21	Please remove the label "QC" from the text and replace by "validation samples", they could be to a great extent identical but not necessarily. Add and distinguish here between LC-MS and ELISA acceptance criteria. **Proposed change (if any):** For the validation of the within-run accuracy, there should be a minimum of five samples per concentration level at LLOQ, low, medium and high QC samples in a single run. The mean accuracy value should be within 15% of the nominal values for the ~~QC~~ *validation* samples, except for the LLOQ which should be within 20% of the nominal value.	It is considered that the use of 'QC' instead of 'validation samples' does not lead to confusion. Therefore the term "QC' has been kept.
199-203	24	Comment: We would like EMA to add criteria for LBA; ULOQ is missing in case of a non-linear concentration-response relationship of the assay. Please refer to CCIII white paper for acceptance criteria.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. The accuracy of the ULOQ should be within 15% of the nominal value.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
199-203	45	1. See above: ULOQ is missing 2. Please remove the label "QC" from the text and replace by "validation samples", they could be to a great extent identical but not necessarily 3. Add and distinguish here between LC-MS and ELISA acceptance criteria	1. The accuracy of the ULOQ should be within 15% of the nominal value. 2. It is considered that the use of 'QC' instead of 'validation samples' does not lead to confusion. Therefore the term "QC' has been kept. 3. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
199- 221	23	The document does not specify whether the QCs prepared for these parameters come from the same prepared batch or whether it is necessary to prepare different QC batches for the between run accuracy and precision.	It is recommended from one batch. Both are acceptable, however this is up to the applicant and will be covered by the SOP. No additional recommendation will be included in the guideline.
200-203	19	There are no criteria for the QC samples per concentration level. Based on these acceptance criteria it may occur that 4 out of 5 QC samples of the same concentration level are outliers. **Proposed change (if any):** The ±15% and ±20% criteria should also be fulfilled for at least 50% of the individual QC samples tested per concentration level as well as for at least 67% of all QCs analysed in the run.	No additional criteria have been included. Furthermore the guideline indicates that determination of accuracy and precision should include all results obtained except those cases where errors are obvious and documented.
201	11	Comment: change "mean accuracy" to "mean bias"	The text has been revised and states now 'mean concentration'.
202 /412	25	Comment: Requirements for accuracy in the validation section of the document is unclear. Accuracy is calculated in %; this should be a number close to 100. Requirement for accuracy is expressed in the document as follows: "The Mean accuracy value should be within 15% of the nominal value of the QC samples. **Proposed change (if any):** Express as in section 5.2: "The accuracy values of the	The text has been revised and states now 'mean concentration'. Regarding the expression of accuracy, it is considered obvious what is understood.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		QC samples should be within ±15% of the nominal values."	
204-208	21	This mixes up inter and intra – assay precision and accuracy; between run with replicates doesn't return inter-assay. Five determinations per run and concentration level seems to be too high to assess variability between the batches. The higher the total number of QC samples the lower the probability to assess a variation. In addition in a normal study batch 2-4 QC samples per concentration level will be analysed. Therefore, to mimic conditions for study sample analysis it would be beneficial to ask for fewer QC samples per concentration level in the two non intra-assay runs. Other statistical approaches as ANOVA are also possible to extract the needed information. **Proposed change (if any):** For the validation of the between-run accuracy at least five determinations per concentration per run at LLOQ, low, medium and high QC samples from three runs analysed on at least two different days should be *preferably* evaluated. *However, other approaches giving the required information are allowed.*	The number of QCs included in an analytical run may also differ between studies. It is acknowledged that the validation procedure may not represent the analysis of study samples in all cases; however this is considered not critical.
204-208	41	Is between-run accuracy the appropriate term?	It is considered appropriate.
204-208	45	1. This mixes up inter and intra – assay precision and accuracy; between run with replicates doesn't return inter-assay, if n = 1 overall number of determination would be only 3 2. Other statistical approaches as ANOVA are also possible to extract the needed information.	For the validation of the between-run accuracy, the results of the QC samples from different runs should be evaluated, and not the mean of each run. An applicant is free to add additional statistical approaches.
204-211	2	Between –run accuracy Comment: • Consider using the already stated greater than 15/20/25% deviations from nominal as failing QCs, and also stating that documented technical errors would also allow exclusion of data.	The question is whether a statistical outlier is true value or a real outlier. Statistics can hide bad performance of the method. Technical errors are acceptable as long as this is clearly

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Additionally a suitable statistical method for determining outliers can also be applied. • Consider not excluding outliers for statistics. An option would be replacing "should" by "can" (as FDA). The results of all runs should be included for validation, except in cases where errors are obvious and documented. • Line 207 & Expand comments on the Between Run (Inter) validation to evaluate P & A differences between equipment, chemists, etc. in the initial validation. • Accuracy & Precision "Within-Run (Intra) and Between Run (Inter)" As harmonization is preferred, provide alignment to terms used in ICH as Intra and Inter (Ref: http://www.ich.org/LOB/media/MEDIA417.pdf Page 9) • LBAs typically evaluate at least 6 runs. Text does not clarify that acceptance criteria are for small molecule only. LBA criteria do not appear until lines 351-352.	documented and in accordance with the SOP. The proposal has been incorporated. This is considered robustness and therefore a method development issue. Within-run and between-run are also indicated other guidance. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
204 – 211	13	Comment: Between- run accuracy should be covered by the determination of intermediate precision and robustness experiments.	The terminology 'intermediate precision' appears not commonly used. Robustness is considered a method development issue.
204-211	24	Comment: We would like EMA to add and distinguish here between chromatography based assays and LBAs acceptance criteria. We would like EMA to consider the ULOQ as well in the case of a non-linear concentration response relationship of the assay. Please clarify how the within run accuracy should be determined. Is it on the collation of all individual values	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. The accuracy of the ULOQ should be within 15% of the nominal value.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		or on the mean for each run? We prefer the terms intra- and inter-assay to align with FDA. An outlier should be determined by a process defined by SOP or the study plan. However, outliers should be included in the overall statistics of the validation. We ask EMA to consider the opportunity to apply other approaches as presented during the Brussels meeting.	For the validation of the between-run accuracy, the results of the QC samples from different runs should be evaluated, and not the mean of each run. The terms are used both. It is considered that between- and within-run is correct. The text has been revised regarding reporting. The guideline prefers standardisation of approaches and is in favour to keep it simple.
204-211	52	Comments: Between-run accuracy should be covered by the determination of intermediate precision and robustness experiments.	The terminology 'intermediate precision' appears not commonly used. Robustness is considered a method development issue.
205	7	Comment: Five determinations per run and concentration level seem to be too high to assess a variability between batches. The higher the total number of QC samples the lower the probability to assess a variation. In addition in a normal study batch 2-4 QC samples per concentration level will be analysed. Therefore, to mimic conditions for study sample analysis it would be beneficial to ask for fewer QC samples per concentration level in the two non intra-assay runs	The number of QCs included in an analytical run may also differ between studies. It is acknowledged that the validation procedure may not represent the analysis of study samples in all cases; however this is considered not critical.
205	26	Why must 3 runs be done on 2 different days instead of 1? If 2 days are better, why not 3 days?	This is a regulatory set criteria.
205-207	6	Comment: Why to analyze each concentration 5-times within run in this kind of test? **Proposed change (if any):** For the validation of the between-run accuracy determination per concentration at LLOQ, low, medium and high QC samples from five runs analysed on at least three different days should be evaluated.	The proposal will also raise questions. 3 runs is a minimum; if more runs are applied this is acceptable.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
205-207	18	Comment: Why to analyze each concentration 5-times within run in this kind of test? **Proposed change (if any):** For the validation of the between-run accuracy determination per concentration at LLOQ, low, medium and high QC samples from five runs analysed on at least three different days should be evaluated.	The proposal will also raise questions. 3 runs is a minimum; if more runs are applied this is acceptable.
205-207	37	Comment: 5 determinations x 3 runs = 15 samples for each concentration lead to larger population to estimate accuracy/precision (n = 15) then proposed by FDA (n = 6). It was shown through years that FDA approach gives acceptable results. Therefore, in our opinion EMA regulations should be similar to current practice accepted by FDA. Moreover, in case of precision significantly higher number of measurements may mask real problems with measurement results quality. No scientific justification, decrease in measurement results quality. **Proposed change (if any):** For the validation of the between-run accuracy at least five determinations per concentration ~~per run~~ at LLOQ, low, medium and high QC samples from at least three runs analysed on ~~at least two~~ different days should be evaluated.	The proposal will also raise questions. 3 runs is a minimum; if more runs are applied this is acceptable. The number of QCs included in an analytical run may also differ between studies. It is acknowledged that the validation procedure may not represent the analysis of study samples in all cases; however this is considered not critical.
206	36	**Between-run accuracy** As long as batches are prepared independently during the day, the requirement for between-run accuracy to be conducted over at least two days is not critical. **Proposed Change:** "from three independently prepared runs should be evaluated"	This request takes into account a kind of robustness; sample analysis will be done at different days and the recommendation in the guideline will mimic in part the study sample analysis.
209-210	41	Comment: It is unclear if the data including the outlier should meet the criteria, or that it is acceptable if the data only meets the criteria excluding the outlier.	The guideline indicates that determination of accuracy and precision should include all results obtained except those cases where errors are obvious and documented.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): Make a more clear description about acceptance of data.	
209-211	6	Comment: This requirement is not consistent with the requirements for the calibration curve, where the outlier should be rejected. We recommend to also allow here the exclusion of the outliers, because otherwise a single outlier caused by operator mistake (e.g. high QC is taken for analysis instead of low QC) would ruin the entire validation. **Proposed change (if any):** Delete this paragraph.	For the calibration curve it will affect the whole run. Regarding the QC, the question is whether a statistical outlier is true value or a real outlier. Statistics can hide bad performance of the method. Technical errors are acceptable as long as this is clearly documented and in accordance with the SOP. The example for the operator mistake will be considered as a normal value (it cannot be checked if this had occurred, only based on outcome).
209-211	19	We agree that statistics without extreme outliers should be allowed. However it would be useful if the guideline suggested a more objective way to decide what values are outliers.	The question is whether a statistical outlier is true value or a real outlier. Statistics can hide bad performance of the method. Technical errors are acceptable as long as this is clearly documented and in accordance with the SOP.
209-211	23	What kind of statistical tests should we use? If a value is statically determined outlier, could the value be excluded from the test and the test validated? Concerning these "outliers" in the between run accuracy, what should be done with the individual values considered as "outliers" and what values should be used in the statistical calculations? The explanations in this paragraph are not very clear. The text states "Reported method validation data….should include all outliers; however, calculations …..excluding values ….should additionally be reported" Is it necessary to include both calculations?	The question is whether a statistical outlier is true value or a real outlier. Statistics can hide bad performance of the method. Technical errors are acceptable as long as this is clearly documented and in accordance with the SOP. Outlier definition is covered by a SOP. This is a case by case evaluation. The text has been revised for clarity.
209-211	27	Comment: Lines 209-211 : "Reported method validation data and the determination of accuracy and precision should include all outliers; however, calculations of	The proposal is considered to elaborate for a guideline. It is acknowledged that different methods exist to detect outliers.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		accuracy and precision excluding values that are statistically determined as outliers should additionally be reported." **Proposed change (if any):** Describe/clarify the method that should be used for outliers exclusion (e.g. USP Pharmacopeial Forum 36 (2) p514 §4 – "OOS results or outliers" referin to <1010> Analytical data – interpretation and treatment)	
209-211	37	Comment: It should be clearly stated that the method is acceptable in case of determined outliers when initial accuracy/precision results are outside acceptance limits and after excluding outliers accuracy/precision results are within acceptance limits. Otherwise, calculations of accuracy and precision excluding outliers are not necessary.	The question is whether a statistical outlier is true value or a real outlier. Statistics can hide bad performance of the method. Technical errors are acceptable as long as this is clearly documented and in accordance with the SOP. Outlier definition should be covered by a SOP. This is a case by case evaluation. The text has been revised for clarity.
209-211	38	Comment: Can you add a definition to Outlier	Outlier: for validation samples identified by an applicable statistical test and no reason can be given for the obtained value.
209-211	39	This requirement is not consistent with the requirements for the calibration curve, where the outlier should be rejected. We recommend to also allow here the exclusion of the outliers, because otherwise a single outlier caused by operator mistake (e.g. high QC is taken for analysis instead of low QC) would ruin the entire validation. **Proposed change (if any):** Delete this paragraph.	For the calibration curve it will affect the whole run. The question is whether a statistical outlier is true value or a real outlier. Statistics can hide bad performance of the method. Technical errors are acceptable as long as this is clearly documented and in accordance with the SOP.
209-211	45	1. In case outliers are excluded then a statistical outlier test should be provided a priori. 2. Distinguish here between LC-MS and ELISA acceptance criteria	1. It is considered to elaborate to include this in the guideline. It is acknowledged that different methods exist to detect outliers. 2. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			assays.
210	49	How do you statistically prove that a value is an outlier. There are many ways of determining this. **Proposed change (if any):** It should be specifically stated what statistical method should be used (eg. Grubbs), so that there is consistency throughout the industry	Outlier definition should be covered by a SOP. This is a case by case evaluation. The proposal is considered to elaborate for a guideline. It is acknowledged that different methods exist to detect outliers.
212	36	**Precision** The term "standard practise" is unclear **Proposed change :** "The statistical method for estimation of the precision should be predefined."	The text has been revised for clarrity.
212-217	2	Comment: • Text does not indicate that criteria are for small molecules. LBA criteria do not appear until lines 345-354. **Proposed change (if any):** • For LBAs, the within-run and the between-run precision should be within ±20%, except for LLOQ which should not exceed 25%.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
212 - 217	24	Comment: We would like EMA to add and distinguish here between chromatography based and LBA based assay acceptance criteria. ULOQ is missing in case of a non-linear concentration-response relationship of the assay	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. The accuracy of the ULOQ should be within 15% of the nominal value.
212-223	45	1. See above comments as for accuracy 2. Give the same level of details as the ones used to determine accuracy 3. Add ELISA acceptance criteria	1/2. The text has been revised. 3. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			separate section has been introduced on ligand binding assays.
213 - 217	10	Comment: According to the guidance precision (within-run and between-run) will be determined with the following samples: LLOQ, low QC, medium QC and high QC. These samples covering only 75% but not the whole calibration range. **Proposed change (if any):** We would propose to add a sample corresponding to the upper limit of quantitation (ULOQ) to cover the calibration range.	The accuracy of the ULOQ should be within 15% of the nominal value.
215	11	Comment: "practise" should be "practice"	The text has been revised and this issue has been deleted.
215	23	**Proposed Change:** Give examples of standard statistical methods for the calculation of precision.	The text has been revised and this issue has been deleted.
215	41	Comment: 'standard practise' for statistical methods is vague. **Proposed change (if any):** Calculations are to be described in the study plan. Re-define or omit the current line.	The text has been revised and this issue has been deleted.
218 - 220	24	Comment: We would like EMA to add LBA acceptance criteria. ULOQ criterion is missing for LBA method and in case of a non-linear concentration-response relationship of the assay, it should comply with LLOQ (see white paper, reference given in line 199)	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
218-220	27	Comment:Line 218-220 – Within-run precision **Proposed change (if any):** Should be aligned with other guidances (e.g. USP Pharmacopeial Forum 36 (2) p512 §2)	The text has been revised for clarity.
221 - 223	24	Comment: We would like EMA to add LBA acceptance criteria. ULOQ criterion is missing for LBA method and in case of a non-linear concentration-response relationship of the assay, it should comply with LLOQ (see CCIII white paper).	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Please clarify how the within run precision should be determined. Is it on the collation of all individual values or on the mean for each run? We prefer the terms intra- and inter-assay to align with FDA.	assays. For the validation of the between-run precision, the results of the QC samples from different runs should be evaluated, and not the mean of each run. The terms are used both. It is considered that between- and within-run is correct.
224	16	Comment: Establishing dilution integrity during pre-study validation may not cover early human trials where first experience with a drug will provide the evidence **Proposed change (if any):** Please permit flexible assessments of dilution integrity with the most rigorous in assays used for later phase development.	It should be known how dilution affects accuracy and precision. This can be covered by partial validation. The text has been revised to clarify this. For LBA dilution integrity is of importance and should be known (in part) on forehand. This is incorporated in the new LBA section.
224-227	23	"Sample shall be diluted with blank matrix", this implies that it is the same species as for the samples. In case of limited matrix (mice, cyno) it might be useful to use another appropriate matrix e.g. human for dilution, given that it is demonstrated that there is no interference. **Proposed change (if any):** … dilution of this sample with an appropriate matrix ….. Comment: Is it necessary to perform this verification during the validation of the method, or can it not be done as a supplementary phase to the validation if and when it is seen during the routine analyses? This would allow the verification of the dilution parameters to be better chosen with respect to the dilutions performed.	Dilution with another matrix is acceptable, as long as it is validated. The text has been revised to clarify this. Using a different matrix for dilution should be evaluated before the start of study sample analysis. During study sample analysis it should be know which matrix is acceptable to dilute the study samples. It may be applicable to do a partial validation, before starting the run with the diluted samples.
224-229	2	Comment: • Establishing dilution integrity during pre-study validation may not cover early human trials where first experience with a drug will provide	The text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		the evidence. • LBA: It is recommended to test dilution integrity (dilution linearity) also with assay buffer. If dilution integrity is shown with buffer, this would save matrix (esp. rare matrix). • It should also be possible to dilute the sample first into assay buffer (the assay's minimum required dilution) and then perform subsequent dilutions in buffer containing the appropriate percentage of matrix. This is sometimes necessary when matrix is costly and/or scarce. It should be ensured that the dilution integrity is performed the same way that samples will be diluted in study. **Proposed change (if any):** • For LBA, the criteria should be 20/25% and samples prepared at concentrations above the ULOQ should measure >ULOQ to demonstrate the absence of a "hook" effect.	For LBA dilution integrity is of importance and should be known (in part) on forehand. This is incorporated in the new LBA section. Dilution with another matrix is acceptable, as long as it is validated. The text has been revised to clarify this. The proposal will be taken into account in the new LBA section
224-229	6	Comment: The dilution integrity should be tested only if there is at least one study samples above ULOQ, i.e. if the dilution was applied in the study. Otherwise it has no sense to study it during method validation. The similar approach as that described under 5.3 (line 430-437) should be applied. **Proposed change (if any):** Insert the sentence: If the analyte concentration in at least one study sample appear to be above the ULOQ, the dilution integrity should be tested.	The text has been revised. It states now 'if applicable'.
224-229	21	For ELISA samples are diluted with assay buffer (at constant matrix concentration) instead of with sample matrix. It is recommended to set up a different experiment for LBA (e.g. serially dilute high	For LBA dilution integrity is of importance and should be known (in part) on forehand. This is incorporated in the new LBA section.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		concentration spikes into 10 individual matrix samples and recommend that at least 60% of all the individuals must have all four dilutions recover within 80-120% of nominal). In addition, a separate experiment to assess a prozone effect should also be conducted, by spiking a sample close to expected Cmax for study. Concentrations diluted within the calibration curve range should be within 20% of nominal and those above the calibration curve range do not quantitate within the assay. Suggest to add a specific paragraph called "Dilution linearity and parallelism" either in section 4.1.7 or 4.4.	
224 - 229	24	Comment: LBA: It is recommended to test dilution integrity (dilution linearity) also with assay buffer. If dilution integrity is shown with buffer, this would save matrix (esp. rare matrix). Is there a justification for at least five determinations per dilution factor? If the maximal DF is validated for LC-MS based assays, all others smaller DF validated are covered. For LBA based assay all used dilution factors must be validated, either during validation or in-study The minimum sample volume that could be diluted should be validated. For LBA, the minimum required dilution (MRD) should be validated.	For LBA dilution integrity is of importance and should be known (in part) on forehand. This is incorporated in the new LBA section. The number of determinations is regulatory set. It is in line with the within run accuracy/precision for validation samples. MRD is incorporated in the new LBA section.
224-229	39	The dilution integrity should be tested only if there is at least one study samples above ULOQ, i.e. if the dilution was applied in the study. Otherwise it has no sense to study it during method validation. The similar approach as that described under 5.3 (line 430-437) should be applied. **Proposed change (if any):** Insert the sentence: If the analyte concentration in at least one study sample appear to be above the ULOQ, the dilution integrity should be tested..	The text has been revised. It states now 'if applicable'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
224-229	45	1. Suggest calling it Dilution Linearity. 2. Could be done during validation or as in-run validation 3. For ELISA samples are diluted with assay buffer (at constant matrix concentration) 4. Five replicates apply only to LC-MS 5. If the maximal DF is validated for LC-MS based assays, all others smaller DF validated are covered. However, For ELISA based assay all used dilution factors must be validated, either during validation or in-study. For LBA only 1. Recommend that a different experiment be set up for LBA (e.g. serially dilute high concentration spikes into 10 individual matrix samples and recommend that at least 60% of all the individuals must have all four dilutions recover within 80-120% of nominal) 2. In addition, a separate experiment to assess a prozone effect should also be conducted, by spiking a sample close to expected Cmax for study, and ensure that concentrations diluted within the calibration curve range quantitates within 20% of nominal and that those above the std curve range do not quantitate within the assay. 3. Add a specific paragraph called "Dilution linearity and parallelism"	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. For LBA dilution integrity is of importance and should be known (in part) on forehand. This is incorporated in the new LBA section.
224-229	47	Comment: Regarding dilution integrity. There is no discussion of dilution linearity, parallelism, and other challenges related to analysis of macromolecules. **Proposed:** As described below for Section 4.4, the guidance appears to lack adequate discussion of immunoassays and/or ligand-binding assays. In this section, dilution linearity and parallelism are not addressed. A distinction	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. Immunoassays are not within the scope of this guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		between dilution linearity and parallelism should be defined, as well as the relevant acceptable use of pooled matrix. Guidance on how dilution linearity should be conducted is recommended, particularly in light of the fact that many laboratories are not consistent with the use of matrices during dilution linearity (and some even assess Prozone effect rather than dilution linearity). Note that several White Papers adequately describe acceptable practices pertaining to the effective use of dilution QC samples and dilution linearity, including the recommendations for addressing bioanalytical analyses that quantitate for samples having concentrations exceeding the validated acceptable dilution and conditions (LBA method must also describe the matrix/buffer composition and make a distinction from Prozone effect analysis). Additionally, it is recommended that the guidance address expectations for samples that are quantitated at a concentration that exceeds the highest validated dilution.	
225	9	For LBAs, validation parameters for dilution linearity and also parallelism should be specified.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
225 - 229	10	Comment: Spiking matrix with the analyte at concentrations > ULOQ results in not measurable concentrations. In this case there will be no basic value to calculate the accuracy after dilution of these samples.	Comparison should laways be made to the nominal concentration, although this concentration is outside the calibration curve range. After dilution the concentration will be inside the calibration curve range and accuarcy and precision can be obtained.
225-229	19	This section does not specify whether dilution integrity is to be tested only in a single run or throughout 3 validation runs (in the same way as accuracy and precision).	One run is considered acceptable; more runs also.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): We suggest for dilution integrity test to be performed on three different occasions and accuracy and precision for the diluted samples to be evaluated within-run and between-run. Acceptance criteria should be the same as for accuracy and precision test.	
225-229	21	Dilution integrity is required (but only once, as this document does not stipulate requirements for multiple batches). We recommend 3 batches of dilution integrity to determine between-run accuracy and precision. **Proposed change (if any):** *Within-run and between run accuracy and precision for* dilution ~~integrity~~ *samples* should be demonstrated by spiking the matrix with an analyte concentration above the ULOQ and dilution of this sample with blank matrix ~~(at least five determinations per dilution factor)~~.	One run is considered acceptable; more runs also.
225-229	22	Comment: Point 4.1.7 Dilution Integrity This section does not specify whether dilution integrity is to be tested only in a single run or throughout three validation runs (in the same way as accuracy and precision). We suggest for dilution integrity test to be performed on three different occasions and accuracy and precision for the diluted samples to be evaluated within-run and between-run. Acceptance criteria should be the same as for accuracy and precision test.	One run is considered acceptable; more runs also.
225-229	26	Line 225. If sample dilutions are expected, dilution integrity should be demonstrated. Line 229. If sample dilutions are not expected, it is not necessary to evaluate dilution integrity in the validation; however, if at a later time dilutions are deemed necessary, dilution integrity must be demonstrated. This information can be added to the validation report.	The text has been revised. It states now 'if applicable'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
225-229 269	48	Comment: - <u>For dilution and stability study</u>: Should we consider the mean accuracy or each individual relating to the acceptance criteria (i.e. deviation < 15% for LC/MS and < 20% for ligand binding assay) If it means mean accuracy, is it acceptable to have more than 1/3 outside of the acceptance criteria? <u>For stability study:</u> No precision acceptance criteria is mentioned. In case of each individual accuracy should be taken into account, the precision is effectively not required. However if the mean accuracy is considered, it is possible to have an accuracy meeting the acceptance criteria despite a CV>30%. Is it acceptable to validate the studied parameter? **Proposed change (if any):** To be clearer in the presentation of the acceptance criteria for the studied parameter (always mention 'individual ' or 'mean' according the case)	The mean accuracy should be taken into account. This does not exclude that borderline outcomes may be questioned.
226	23	Post extraction dilution remains feasible?	Post extraction dilution may be critical. Post extration dilution should ensure that the extraction recovery is the same. Moreover, there may be a poblem with matrix effects.
226 - 229	23	This paragraph does not identify at what concentration (one or several) these tests must be performed; mid range, LLOQ, ULOQ, Since you discuss the precision of dilution should these tests be performed in one single run or repeated several times. That is intra or inter run precision.	The guideline indicates clearly that it should be a sample spiked at a concentration above the ULOQ.
227	8	Comment: The dilution integrity evaluation typically evaluates that a very high concentration can accurately read after dilution into the validated range. This is not necessarily dependant on dilution factor. Therefore, is it acceptable to test only 1 dilution factor? Many factors are implied by the guidance. How many? Which	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		factors?	
228-229	19	What should we do if during sample analysis it is found that a higher dilution factor is required than that tested during method validation? Should dilution integrity be re-validated or is it acceptable to investigate a higher dilution factor in-process during sample analysis? **Proposed change (if any):** We suggest to include that in such a case an additional dilution factor can be validated in-process during sample analysis.	Dilution integrity can be covered in a partial validation.
228-229	22	Comment: Point 4.1.7 Dilution Integrity With the current wording, it is unclear what should be done in a situation where during sample analysis it is found that a higher dilution factor is required than that tested during method validation. Should dilution integrity be re-validated or is it acceptable to investigate a higher dilution factor in-process during sample analysis? **Proposed change (if any):** We suggest to include that in such a case an additional dilution factor can be validated in-process during sample analysis.	Dilution integrity can be covered in a partial validation.
228-229	37	Comment: As it is assumed that calibration range is sufficient to determine study samples concentration, it should be clearly stated that dilution integrity is not necessary part of method validation, it can be performed after analysis of study samples and is not required until samples are diluted during the study.	The text has been revised. It states now 'if applicable'. For LBA dilution integrity is of importance and should be known (in part) on forehand. This is incorporated in the new LBA section.
229	23	Should we test in range dilution during validation to be used when insufficient sample volume is provided? If yes, what concentration levels should be tested?	For chromatographic methods, the applied dilution factor should be covered. For LBA dilution integrity is of importance and should be known (in part) on forehand. This is incorporated in the new LBA section.
230	20	Comment: Matrix effects (ME) is an extremely important topic in LC-MS/MS methods that often is not suitably	Comment is noted

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		addressed during method development (MD). In practice, the standard approach for evaluating ME (based on overall %CV) is not extremely effective in identifying significant ME. I have seen > 5 assays that met these validation criteria but failed during sample analysis due to ME. Personally, I believe that we need to implement a new MD paradigm where we not only build in quality and minimize ME through sound scientific strategies, but we should also try to make methods fail during MD rather than just passing set criteria. However, until this time comes, it is recommended that more stringent criteria be included for routine ME evaluation during MD **Proposed change (if any):** See specifics below	
230 - 233	24	Comment: We think that standardization of the degree of haemolysis and lipidemia of a sample is difficult to achieve and therefore it would be hard to address the impact. This should be investigated based on scientific judgment and during method development. A tiered approach is recommended. In cases where isotope-labeled internal standards are used, the scientifically accepted position is that these internal standards will compensate for potential matrix effects, the use of one source of matrix would be sufficient. If non-isotopically IS is used, at least 6 sources of matrix is recommended. Please harmonize with the FDA white paper. Matrix effects play a very significant role in binding assays. The definition of a matrix effect needs to be expanded to binding assays, and also the methodologies to assess these. Please add definitions/criteria also for LBA	The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic. Other papers indicate that this did not compensate for matrix effect (due to a shift in retention time). It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
230-256	2	Comment: • It is rather difficult to obtain haemolyzed, hyperlipidaemic matrix, as well as matrix from special populations as renally or hepatically	The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		impaired populations. It may occur that the validation activities may considerably delay projects due to the need to find or even generate those special matrices. Only the testing of haemolyzed plasma may be acceptable if at all. All other cases of special matrix cases are simply not feasible within a reasonable time frame. Consider rephrasing or removing the paragraph.	haemolysed and hyperlipaemic.
		• No standard level of haemolyzed, hyperlipidaemic matrix is provided and the laboratory is left to define what is appropriate.	This is agreed. It is difficult to give a standard level.
		• When implementing an assay for study sample analysis, bioanalytical labs do not routinely receive clinical chemistry reports on each subject, such that it would be possible to ensure that the levels of hemolysis or lipidaemia tested in validation cover those of the study samples and it would not be practical to implement such practices.	This is agreed.
		• It's costly and time consuming to investigate the excipients and their metabolites that were previously obtained from subjects or animals. Moreover, it adds little quality to the method compared to the efforts taken. Finally, the isotope-labeled IS should compensate the effects generated by excipients and their metabolites.	It is recognized that it can be time consuming, however it adds clearly to the quality. Isotope labelled IS can also not compensate always for this
		• 6 batches for human matrices are ok. At least 2 batches for animals are sufficient as for specificity.	Although laboratory animal are standardized in specie, housing, food, still the animals are individuals. There is no demonstration that it would not be a problem for animals.
		• LBA: Highlight in the section title that this paragraph only applies to chromatographic / mass spectrometric methods.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		• Line 241: MF determination is an emerging science and MF can be determined for on-line extractions also. Mass spectrometer vendors are now creating software for such a determination. There is no good substitute for the MFs. MFs provide a very useful parameter for assessing the assay performance in multiple matrices, e.g., multiple matrices of unknown test samples. The requirement for determination of MF for online extractions should, however, be relaxed and replaced by equivalent experiments until the science is mature. • Lines 241-249: For determining the concentrations, one would require calibration standards to be prepared in one of the lots. Which lot should one consider as the calibration backbone? Concentration determination suggested here is equivalent to determination of precision and accuracy of low QC in 6 different lots. Instead of determining the concentrations; assessment of the peak areas of analyte, IS and their ratios would be better parameters to determine for online extraction to determine the MF and IS normalized MF.	recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. The comment is noted. A lot should always be used for such determinations. This can be the same or a different lot. Peak areas should be reported in addition to the concentration. From those areas it can be observed if there is an effect.
230-256	26	Line 232. ..6 lots of matrix are recommended, but how many with these special conditions, e.g., haemolyzed, hyperlipaemic, etc.? Line 233. In theory a good idea, but leads to a question: How are these results used? If 10%, but not 5%, haemolyzed sample shows an effect, must an assay be developed to determine the extent of haemolysis in each sample? Must separate calibration standards and QCs be used for 10% haemolyzed matrix, etc.? Is any sample showing haemolysis not assayed? etc.	The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic. It is recognized that a clear definition on extent of haemolysed cannot be given.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Line 234. Use peak areas, not ratios. Perform with pre-extraction spikes, post-extraction spikes and solvent standards to get True Recovery, Combined Recovery, and Matrix Suppression/Enhancement. Line 252. In theory this is a good idea, but potentially it is difficult to implement in practice. Determining concentrations of excipients to evaluate is problematic, as these should be relevant to excipient concentrations in the subject after dosing – which are generally unknown. **Proposed change (if any):** Clarify guidance appropriately	It is considered that the text is sufficient clear. In addition, recovery is not requested. The concentration of the excipient in the matrix is a result of the treatment. As stated, the matrix should be obtained from treated subjects or animals. The matrix should not be spiked, as this does not represent the real matrix containing the excipient and it metabolic compounds.
230-256	45	1. This section is particularly important for ELISA, less for LC-MS assays when an isotopically labeled internal standard is used. The definition of a matrix effect needs to be expanded to binding assays, and also the methodologies to assess these. Please add definitions/criteria also for ELISA. 2. For LC/MS/MS assays stable isotope labeled internal standards provide an excellent opportunity to make the assay very rugged. Use of such ISs should be highly recommended. Having performed MFs for analyte and ISs for the last 10 years we have data to prove that when stable isotope labeled ISs are used, it is not necessary to determine the MFs in 6 different lots of matrix. The IS normalized MFs in such cases is always close to 1. The requirement for 6 lots in such cases should not be required. This will encourage the sponsors to make an extra effort to synthesize the stable isotope labeled ISs and make the assay more rugged. 3. In addition, for animals there is no change in matrix used during validation and study, in humans there is no source of matrix available which has a close relationship to the actual	1. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. 2. It is recognized that labelled IS can cover this but not in all cases. On forehand it is not known which cases these are. 3. Comment noticed. 4. Yes. 5. A separate section has been introduced on ligand binding assays. 6. This is agreed. 7. The section has been revised and indicates that MF should be evaluated at a low level (maximum 3 times the LLOQ) and high level of concentration (close to the ULOQ) 8. Both formulas are the same. It is a different way of writing the formula.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		study sample population (common practice) 4. Haemolysed and hyperlipidaemic applies for blood, plasma and serum only. 5. MF can be determined e.g. in comparison of dilution linearity of analyte X in serum and analyte X in serum free buffer. (ELISA) 6. If it is scientifically warranted then different subtypes of matrix should be investigated 7. We suggest to study MF at 3 different concentrations (LC-MS methods) corresponding to low QC, med QC and hi QC 8. Normalized MF = ratio of analyte/IS added to matrix after extraction over analyte/IS ratio in matrix free solution	
231	9	Matrix effect should also be tested for ligand binding assays using appropriate experimental design.	It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays.
231	16	Comment: Testing of hemolysis and hyperlipidemia as described create two potential problems: 1) No standard level of either is provided and the laboratory is left to define what is appropriate. 2) When implementing an assay for study sample analysis, bioanalytical labs do not routinely receive clinical chemistry reports on each subject, such that it would not be possible to ensure that the levels tested in validation cover those of the study samples and it would not be practical to implement such practices. **Proposed change (if any):** Please provide a recommended level of hemolysis and hyperlipidemia, and the assurance that assessment of individual study samples will not be needed during study sample analysis.	1. It is agreed that a clear definition on extent of haemolysed cannot be given. 2. This is agreed. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
231 - 232	20	Comment: The 6 sources of plasma used for selectivity	The text has been revised, clearly indicating that pooled

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		and matrix effect evaluation should be obtained from individual donors or animals rather than lots pooled from multiple donors or sources. Significant Matrix Effects (ME) in a single lot of plasma may be sufficiently diluted in pooled matrix to a point where significant ME will not be identified. **Proposed change (if any):** Recommend clarifying that the plasma lots should be obtained from individual donors or animals	matrix should not be used.
231 – 232	23	What is the consequence if the CV is more than 15%? Does this render the method invalid? The matrix effect only applies to the mass spectrometric methods? It is not required for classic HPLC chromatography? It is unclear if we should use 6 lots of each of normal, haemolysed and hyperlipidaemic matrix. Could you indicate what degree of haemolysed and hyperlipidaemic matrix should be tested. How should this be documented? The use of 6 lots of matrix should only be used if possible. Should matrix effect be tested on hyperlipidaemic plasma for any validation or could it be related to some study only? In this case, how should we document this and what kind of study is concerned? Should this be tested in every validation study – or is it sufficient to demonstrate it only if such samples are expected/observed? Could the test for special populations be implemented during the clinical/toxic study by spiking pre-dose study samples. Perhaps this requirement should be put in as an option. **Proposed Change:** Replace "Matrix effects should be	A CV of more than 15% indicates that the method is not reliable/accurate. Therefore the method should be changed (sample preparation and/or analysis for instance). MF is only applicable to LC-MS methods. It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		investigated when using mass spectrometric methods using at least 6 lots of normal, haemolysed and hyperlipidaemic matrices…". The limitations concerning the renal and hepatic deficiency population could also apply to the cases of the different matrices (hyperlipaemic etc., In fact this restriction "if necessary" in place of "if necessary" could include the whole paragraphe and not just the end of the sentence.	
231 – 232	35	Comment: It is not clear how many lots of haemolysed and hyperlipidaemic are to be taken out of the 6.	The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
231-232	40	Comment: The purpose of the matrix effect is to ensure that the method is not affected by the matrixes from different donors. The key objective is not to assess the effect of haemolysis or hyperlipidaemic. **Proposed change (if any):** … using at least 6 lots of matrixes from different donors and if applicable, haemolysed, hyerlipidaemic and sample matrixes from special populations …	The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
231-233	8	Comment: Stating that haemolysed and hyperlipidaemic plasma is required for the matrix effect evaluation, is too prescriptive. These should be tested when scientifically necessary. If method is robust, especially with a stable-labeled IS, these are typically unnecessary evaluations. **Proposed change (if any):** Remove requirement.	The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
231-233	8	Comment: Is it necessary to demonstrate the absence of matrix effect on the matrix used for the precision and accuracy batches? The lot used to spike these batches might be the deviating lot, and yet, this would not be evident without requiring this test.	The choice of lots to be tested is free. As long as it is tested.
231 - 233	10	Comment: Very often such samples will not be to obtain	The text has been revised. 6 lots for normal matrix should be

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		easily (exception: haemolytic). If the analytical method will be used to determine the analyte in samples of healthy volunteers it is not necessary to introduce samples from volunteers or patients with different diseases into the validation.	evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
231-233	14	Comment: It is not clear whether this paragraph also refers to animal plasma matrices. In cases where isotope-labelled internal standards are used, the scientifically accepted position is that these internal standards will compensate for potential matrix effects in all those cases where for the stable-labelled molecule an isotope effect can be excluded. Moreover, it is much easier to exclude haemolysed or hyperlipidaemic samples from PK evaluation in case implausible results were obtained for these samples as to conduct complex matrix effect investigations. In the guideline text we propose to avoid giving elaborated recommendations for the conduct of matrix effect investigations. Matrix effects should be investigated based on sound scientific judgement. **Proposed change:** A more appropriate wording would be: "*Matrix effects should be investigated when using mass spectrometric methods, using at least 6 lots of individual matrix. For human pk studies sample matrix from special populations, such as renally or hepatically impaired populations or lots of haemolysed or hyperlipidaemic plasma should be considered if matrix effects are expected*."	The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic. It is also applicable for animal plasma matrices. It is recognized that labelled IS can cover this but not in all cases. On forehand it is not known which cases these are. Exclusion for PK results would not be acceptable. It should be evaluated to proof this.
231-233	15	Please distinguish between the validation efforts for samples from animal studies and special populations for human studies. **Proposed change:** Matrix effects should be investigated in more detail for	It is also applicable for animal plasma matrices. It is recognized that labelled IS can cover this but not in all cases. On forehand it is not known which cases these are.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		mass spectrometric methods, when samples from special populations, such as renally or hepatically impaired populations or lots of haemolysed or hyperlipidaemic plasma samples will be analysed, especially in cases where no stable isotope labelled internal standard will be used.	The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
231-233	19	Availability of haemolysed and hyperlipidaemic matrices can be problematic. Additionally, what degree of haemolysed and hyperlipidaemic matrices should be used? We do not suggest to test these kind of matrices in every validation. **Proposed change (if any):** These requirements should be limited to special studies (for example hyperlipidaemic matrix should not be tested for animals).	It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
231-233	22	Comment: Point 4.1.8 Matrix Effect The availability of haemolysed and hyperlipidaemic matrices can be problematic. Additionally, the degree of haemolysed and hyperlipidaemic matrices to be used needs to be specified. Studying the effect of hyperlipidemic matrices will certainly prove challenging. Therefore, we do not recommend to introduce the systematic testing of these matrices in every validation. **Proposed change (if any):** These requirements should be limited to special studies (e.g. not applicable for studies on animals).	It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
231 233	31	Comment: Matrix effects should be investigated when using mass spectrometric methods, using at least 6 lots of matrix including haemolysed, hyperlipidaemic and if applicable, sample matrix from special populations, such as renally or hepatically impaired populations. There are no commercial suppliers of haemolysed and hyperlipidaemic matrixes. **Proposed change:**	It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		It is suggested to specify hyperlipdeamic matrix and the degree of haemolysis in order to have standardized matrixes.	
231-233	37	Comment: It is wider description than FDA, but it is a reasonable change, which leads to improvement of measurement results quality.	Comment noted.
231-233	41	Comment: 1) If haemolysed and hyperlipidaemic plasma is to be included, it should be defined what is considered haemolysed or hyperlipidaemic plasma. 2) If the developed method is sufficiently rugged, this should be of no concern and may not need to be demonstrated. 3) Special populations should be included. It is often not clear to which populations the drug will be applied to; are Caucasians and Asian populations different in this concept? Are males and females different in this respect? **Proposed change (if any):** 1) An industry standard should be defined 2) Possibly exclude from validation, but stress the importance and make it part of method development. It is difficult to obtain reliable hyperlipidaemic plasma. 3) Typically this cannot be validated upfront, but should be part of an investigation once such a study is executed. Typically, studies start in healthy (male) volunteers, and drug development is often halted prior to reaching 'special population studies'. It would be overkill to always include this in validations.	1. It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic. 2. This can only be demonstrated by testing. 3. The text has been revised, mentioning special populations.
231-233	43	(The recommended modifications are in bold.) Matrix effect should be investigated using LC/MS methods especially in cases when the analyte and internal standard possess very different proton (electron) affinity or if the composition of components eluted from the column into the ion source is changing significantly in time. In cases that both the analyte and the internal standard are strong Bronsted bases (in the	The proposed text is considered to detailed for a guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		positive ion mode MS) or strong Bronsted acids (in the negative ion mode MS) matrix effect might be neglected and needs not to be investigated. Generally, matrix effect needs not to be investigated if both analyte and the internal standard possess identical (or very similar) physico-chemical properties. This is generally valid without any limitation in cases when internal standards are isotopically labelled with the next stable isotopes: ^{13}C, ^{15}N, ^{18}O. In case of usage of the isotope ^{2}H for the labelling, the number and position of ^{2}H atoms in the molecule should not be grater than 4 for to exclude the higher possible isotopic kinetic effect which might influence the relative ions yield. The low isotopic effect here should be controlled and ensured by keeping of the identical retention time of both analyte and the internal standard during the analyses. Moreover, the hydrogen exchange reactions must be avoided in the whole sample processing and sample measurement process. In cases when non-isotopically labelled internal standards are used a special attention should be paid to the chromatographic separation conditions of both analyte and the internal standard. The difference of the retention times of both the analyte and the internal standard should be minimised as much as possible for to ensure the identical ionization conditions for both the analyte and the internal standard. At least 6 lots of matrix including haemolysed, hyperlipidaemic and if applicable, sample matrix from special populations, such as renally or hepatically impaired populations. For each analyte and the internal standard, the matrix factor (MF) should be calculated in each lot of matrix, by calculating the ratio of the peak area in the presence of matrix (measured by analysing extracts of blank matrix spiked with analyte at a concentration of maximum 3 times the LLOQ after extraction), to the peak area in	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		absence of matrix (pure solution of the analyte). The IS normalised MF should also be calculated by dividing the MF of the analyte by the MF of the IS. The CV of the IS-normalised MF calculated from the 6 batches of matrix should not be greater than 15 %. If the CV of the IS-normalised MF is greater than 15%, another type of the internal standard must be used and a new IS-normalised MF must be evaluated. If this approach cannot be used, for instance in the case of on-line sample preparation, the variability of the response from batch to batch should be assessed by analysing at least 6 batches of matrix in triplicate, spiked at a concentration of a maximum of 3 times the LLOQ. The validation report should include the peak areas of the analyte and of the IS and the calculated concentration for each individual sample. The overall CV calculated for the concentration should not be greater than 15 %. The mean concentration should be within 15 % of the nominal concentration. The mean concentration should also be reported for each individual batch of matrix; a deviation of this mean from the nominal concentration of more than 20 % in any individual batch of matrix should lead to additional investigations. If the matrix is difficult to obtain, less than 6 different batches of matrix may be used, but this should be justified. However, matrix effects should still be investigated. If a formulation for injection to be administered to the subjects or animals contains excipients known to be responsible for matrix effects, for instance polyethylene glycol or polysorbate, matrix effects should be studied with matrix containing these excipients, in addition to blank matrix. The matrix used for this evaluation should be obtained from subjects or animals administered the excipient, unless it has been demonstrated that the excipient is not metabolised or transformed *in-vivo.*	
231, 239, 250	23	The term “lots of matrix” and batches of matrix” could be harmonized which would help in preventing confusion	The text has been revised and states now ‘lots’.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		in the interpretation of this guideline.	
231, 253,254	17	Comment: "....matrix effects should be studied with matrix containing these excipients..." Proposal: It should be pointed out that this has not to be done with 6 different batches of matrix	Excipient interaction is a different issue. 1 lot is considered sufficient.
231-256	21	While scientifically a sound practice, investigation of matrix effects will place undue burden on the process of method development. Acquisition of the various matrices listed is not trivial. Consider that incurred sample reanalysis of actual patient samples will uncover issues with inhomogeneity and the root cause could then be subsequently investigated. Similarly, comparison of internal standard responses of study samples versus standards/QCs analyte response will uncover the matrix effects. The more important consideration should be that adequate sample preparation and/or chromatographic resolution should be incorporated into a method to minimize the potential for matrix effects. Because of limited feasibility to obtain all recommended matrices and other practicalities the matrix effect assessment described is cumbersome. Requirements for matrix effects of heamolysed, hyperlipidaemic, etc matrix should not be studied in the validation section, but rather in the study analyses section. We recommend that this is not a requirement for validation and allow these to be addressed with study specific samples within the conduct of the study (i.e., samples from control group animals, predose samples from special populations), if needed. **Proposed change (if any):** Please replace text from line 230 through 256 by the following: **4.1.8 Matrix effect**	The proposed change is considered not to improve the current text. Moreover, the text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Matrix effects should be investigated when using mass spectrometric methods because the presence of unmonitored, co-eluting compounds from the matrix may affect the detection of analytes. Quantitative measurement of a matrix effect can be performed by determining the Matrix Factor (MF) which is defined as the ratio of the analyte peak response in the presence of matrix ions to the analyte peak response in the absence of matrix ions (i.e., pure solution). Peak response is defined as the peak area, peak height, peak area ratio, or peak height ratio of chromatographic peaks. The determination of matrix factors should be performed by using at least 6 independent sources of matrix. If the matrix is rare and hard to obtain, the requirement for assessing variability of matrix factors in 6 lots can be waived. The variability in matrix factors, as measured by the coefficient of variation (CV), should be less than 15 %. An MF of 1 signifies no matrix effects. An absolute MF (or IS-normalized MF) of about 1 is not necessary for a reliable bioanalytical assay. However, highly variable MF in individual subjects would be a cause for the lack of reproducibility of analysis and requires investigation. Stable isotope labeled internal standards help by normalizing MF to a theoretical value of 1. If this approach cannot be used, for instance in the case of on-line sample preparation, the variability of the response from batch to batch can be assessed from the overall CV of calculated analyte concentrations of standards and/or QCs. The CV should not be greater than 15 %.	
231-256	22	Comment: Point 4.1.8 Matrix Effect The current wording of this section presents 2 options (option 1: lines 234-240 vs. option 2: lines 241-249) to establish the matrix effect however, the actual benefit of	Line 241 clearly indicates that if option 1 cannot be used, option 2 can be followed.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		option 2 is not clear. The EGA would recommend to specifically include the following two parameters as being crucial in the evaluation of the matrix effect during bioanalytical method development and validation: 1. Ion suppression (Matrix Factor: ion enhancement or suppression) 2. Variability of the effect between matrices (to establish the degree of variability of the matrices effect, and consequently their effect of the method's precision) The EGA believes these two aspects are needed to demonstrate that quantification is not affected by the matrix effect. In addition, we would appreciate to have clarification as to whether experiments for assessment of matrix effect according to lines 234-240 (option 1) have to be performed in triplicate (as mentioned for option 2 in line 243)?	For option 1 as well as for option 2, it is left to the applicant to which extent (i.e. number of replicates per lot) the matrix effect is evaluated.
232	4	Comment: We consider the statement of "haemolysed, hyperlipidaemic" is too vague. **Proposed change (if any):** Insert: "greater or equal to a haemocrit of 5%(v/v)" before "haemolysed" and "plasma from hyperlipidaemic subjects". This should limit people just choosing a, 'pink plasma' or a 'fatty looking plasma' for their validation studies.	It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
232	7	Comment: Matrix effects should not need to be investigated in haemolysed and hyperlipidaemic matrix as well as in matrices from special population. The degree of haemolysis or hyperlipidaemia is a subjective assessment. Of course, haemolytic plasma can be	It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		simulated by adding whole blood to plasma. But how haemolytic is the simulated matrix with respect to the study matrix? The use of a stable-labelled internal standard usually compensates for any difference between spiked matrix samples and study samples.	It is recognized that labelled IS can cover this but not in all cases. On forehand it is not known which cases these are.
232	8	Comment: If haemolysed and hyperlipidaemic are required, at what levels? Also, it is very difficult to evaluate if a sample is hyperlipidaemic, so what is the value of this test?	It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
232	23	We are of the opinion, that matrix effects in haemolysed and/or hyperlipidaemic matrix is problematical: What degree of haemolysis or hyperlipidaemic and how to document this?	It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
232	36	**Matrix effects** **Comments** The implication in the guidance wording is that haemolysed and hyperlipidemic matrix are always to be assessed. This is only applicable to plasma matrix. Other matrices are also assayed in clinical trials**.** **Proposed changes :** Matrix effects should be evaluated as needed for applicable sample conditions (e.g., haemolysed, hyperlipidemic) or special populations (e.g., samples from renal or hepatic impairment studies). The effects and the acceptance criteria to be evaluated should be documented prior to the experiments being conducted	It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic.
232	49	Matrix Effect: Whether haemolysis significantly increases the normal plasma matrix effect or not is debatable. However a major consideration regarding haemolysed samples has been overlooked. When a drug has a large volume of distribution or accumulates specifically in erythrocytes, haemolysis will seriously affect the concentration measured in that particular plasma sample. From our experience this effect may be seen as	The comment is agreed, however this is a distribution issue and not an analytical issue.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		an abnormal spike in the expected kinetic profile and can even appear to be Cmax even though it is sometimes found during the elimination phase.	
232,233	47	Comment: Regarding haemolysed and hyperlipidaemic matrix. The document states that haemolysed and hyperlipidaemic matrix effects should be investigated, but there is no guidance on how this should be done, or the justification for performing this investigation since quantification of haemolysis and lipemic samples is neither easily defined nor currently practiced in industry. Although many laboratories monitor for haemolysis (including more CROs so as to have relevant data to determine if variations in results and/or response signal are a result of haemolysis), the effects of this phenomenon and that of lipemic effects should not affect the bioanalytical method if it is sufficiently robust (whether a chromatographic or LBA). For example, when a method utilizes a stable-labeled internal standard, an LC/MS method that is sufficiently robust will not be impacted by matrix effects resulting from haemolyzed or lipemic samples. If a method is impacted by these samples, then the method is simply not sufficiently developed and/or validated. This situation should require the same attention as an ISR failure, for example an investigation and potentially revalidation. **Proposed:** Recommend that investigation of haemolysis and lipemic effects not be included in the guidance, since it is a bioanalytical issue rather than a method issue. We feel that it is prudent to indicate that the bioanalytical data should never be ignored, and therefore it is still advisable to monitor haemolysis. In short, we recommend tracking haemolysis, but it is not necessary to prescribe guidelines for validation since this phenomenon is usually investigated during method development.	It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic. It is recognized that labelled IS can cover this but not in all cases. On forehand it is not known which cases these are.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
232-233	50	Comment: Matrix effect should be investigated when using mass spectrometric methods, using at least six lots of matrix including haemolysed and hyperlipidaemic and if applicable sample matrix from special population such as renally or hepatically impaired populations. An appropriate procedure for identification (e.g. cobas test for triglycerides-lipids or haemolysis indicator, plasma spiked with 2-5% of whole blood, for haemolysed samples) and extent of haemolysis and hyperlipidaemic samples, and minimum no. of haemolysed and hyperlipidaemic plasma lots under investigation should be should be incorporated in section 4.1.8 (Matrix Effect).	It is recognized that a clear definition on extent of haemolysed cannot be given. The text has been revised. 6 lots for normal matrix should be evaluated. The text indicates now 'recommended' for haemolysed and hyperlipaemic. Inclusion of an identification procedure is considered beyond the scope of this section.
233	23	This is fairly easy for humans but it should be noted that it is not be required for other species. **Proposed Change:** This requirement should therefore be limited to clinical studies.	The evaluation of matrix effects is not limited to clinical studies.
234-237	23	There is now no discussion on recovery studies although this section comes close to it for MS. The difference however is that there is no extraction tested as in recovery testing.	Recovery testing is considered a development issue and therefore not included in this guideline.
234-238	5	Comment 1: For the investigation of the matrix effect a concentration of maximum 3 times LLOQ should be used. The variability of low concentrations is often higher than for e.g. QC-M. Therefore, it would be helpful to perform the investigation of matrix effect with a concentration of QC-M which is "sensitive" for matrix effect but has lower variability than e.g. 3 times LLOQ. Comment 2: It should be mentioned that the IS normalised MF can also be obtained by dividing the peak area ratio (analyte/IS) in presence of matrix by the peak area ratio (analyte/IS) in absence of matrix.	1. The section has been revised and indicates that MF should be evaluated at a low level (maximum 3 times the LLOQ) and high level of concentration (close to the ULOQ) 2. Both formulas are the same. It is a different way of writing the formula.
234 - 238	24	Comment: There should be a differentiation between assays using stable isotope labelled standards and those	This is not agreed, as the obtained CV is method independent.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		which do not. With isotopically labelled IS, normalized MF should be close to 1, with non isotopically labelled IS, CV measured on normalized MF should be calculated and be ≤15%.	
234-238	37	Comment: The introduction of a new term *matrix factor* seems not to be necessary. We suggest to use terms: absolute matrix effect (post extraction spiked vs. pure solution area ratio) and relative matrix effect (CV, IS-normalized) as suggested in literature [B.K. Matuszewski et al., Anal. Chem. 75 (2003) 3019-3030] The nature of matrix effect, i.e. ionisation suppression or enhancement, requires in our opinion investigation of both low and high level QC. As absolute matrix effect does not influence method reliability it may be unnecessary to investigate it. In case it is still recommended (but rather not required) to investigate absolute matrix effect, it should be clearly stated that: a) absolute matrix effect of 85-115% is interpreted as absence of absolute matrix effect b) presence of absolute matrix effect is acceptable c) relative matrix effect should not exceed 15% (CV ≤15%)	Matrix factor is not a new term. It is for instance a term used in the White paper terminology. a. The text is considered sufficient clear that the CV should not be greater than 15%. b. As long as it is consistent, yes. The CV should be within 15%. c. The text is considered sufficient clear that the CV should not be greater than 15%.
234-240	6	Comment: Taking into consideration that recovery testing is not included in the proposed guideline, the matrix-effect test at analyte concentration near LLOQ is only relative matrix effect test and as such, it does not assess a matrix effect in the whole range of the calibration curve.	The recovery vs. matrix effect relation suggested here is not understood. The section has been revised and indicates that MF should be evaluated at a low level (maximum 3 times the LLOQ) and high level of concentration (close to the ULOQ)
234-240	18	Comment: Taking into consideration that recovery testing is not included in the proposed guideline, the matrix-effect test at analyte concentration near LLOQ is only relative matrix effect test and as such, it does not assess a matrix effect in the whole range of the calibration curve.	The recovery vs. matrix effect relation suggested here is not understood. The section has been revised and indicates that MF should be evaluated at a low level (maximum 3 times the LLOQ) and high level of concentration (close to the ULOQ).

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
234-240	19	How many replicates should be tested from each lot of matrix? Please include whether this test is to be perform using single samples or duplicates/triplicates. Acceptance criteria should be reviewed when using duplicates/ triplicates.	For option 1 as well as for option 2, it is left to the applicant to which extent (i.e. number of replicates per lot) the matrix effect is evaluated.
234 - 240	20	Comment: The IS normalized MF approach may be beneficial in methods with stable label IS that coelutes with analyte, but may not be adequate for methods with analog IS that do not coelute with analyte. Co-eluting stable label IS materials have also been shown to cause ME for analyte in some circumstances. Since ME are not necessarily reproducible from sample-to-sample, in practice, the IS normalized MF approach in 6 lots of plasma may not truly reflect the extent of ME observed across the many individual lots of plasma evaluated in clinical studies. Criteria for individual MF should also be included to identify analytes and/or IS where significant ME are present **Proposed change (if any):** Recommend evaluation of ME arising from the presence of IS and recommend adding an additional requirement that IS normalized MF for 5 of the 6 lots evaluated must fall within the range of 0.85 to 1.15. Also recommend performing additional ME evaluations (such as post-column infusion and/or phospholipid transition experiments) if individual MF for analyte or IS in the individual matrix lots falls outside of the range of 0.85 to 1.15	The text is considered sufficient clear that the CV should not be greater than 15%. Additional evaluation can be added, but this is not requested.
234-240	40	Comment: The proposed matrix effect is assessed by CV% of the matrix factor (MF), which is the level of plasma suppression or enhancement. Basically, this approach can ensure that absolute plasma suppression or enhancement CV from 6 lots is below 15%. But this does not mean there is no matrix effect to the analysis. I include a table: "Illustration of matrix effect based on proposed guideline" to demonstrate my point. (My apology, the table cannot fit in this form). As you can see, although the IS normalised CV <15%, 50% sample	The text is considered sufficient clear that the CV should not be greater than 15%. However this does not exclude that study outcomes may be questioned in case of borderline outcomes. The section has been revised and indicates that MF should be evaluated at a low level (maximum 3 times the LLOQ) and high level of concentration (close to the ULOQ).

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		exceeds the acceptance criteria. Another point is that the suggested concentration is 3 times LLOQ which covers the low concentration only. However, when there is a matrix effect enhancement, the impact is higher at high concentrations. **Proposed change (if any):** Matrix effect assessment should be performed by extracting low and high QC samples spiked in six different lots of matrix. The samples for the matrix effect assessment should be extracted along with a calibration curve. The evaluation will be bases on the back-calculated concentrations from the calibration curve. Acceptance Criteria: At least 67% of matrix effect QC samples at each level must have a % R.E. within ±15.0% and all 6 matrix lots must be acceptable. Note: I include a sample of matrix effects from our report for your information.	
234-249	8	Comment: The guidance appears to state that the matrix factor is the only allowable method to determine the matrix effect. Other, equivalent options should be allowed.	The method recommended is considered most straight forward. Other methods may be acceptable on exceptional basis, as long as this is scientifically justified.
234-240	39	The easier approach which would yield the same results would be to spike 6 lots of matrix with a mixture analyte/IS. The CV of the analyte/IS peak area ratio should be less than 15%.	The proposed approach seems more or less a similar approach than option 2 in the guideline. The guideline is considered sufficient clear.
236	9	Need to clarify comment "blank matrix extract spiked ..." If blank matrix is spiked with analyte, extracted and compared to pure solution, there is both a recovery and matrix effect. Spike needs to be into an extract of blank matrix to assess matrix effect.	The text has been revised to cover this issue.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
239,231	9	The words "lot" and “batch” are used interchangeably in the text **Proposed change (if any):** Suggest to use one term consistently; term "lots" is preferred.	The text has been revised and states now ‘lots’.
241	37	Comment: Standard line slopes used to measure relative matrix effect in such a cases may be an interesting option [see: B.K.Matuszewski, J. Chromatogr. B 830 (2006) 293-300].	Other methods may be acceptable on exceptional basis, as long as this is scientifically justified.
241	45	In reference to line 241, we would like to suggest that sponsors should be encourage to perform MF experiments even when the method is an on-line extraction. We have routinely determined MFs for on-line extractions and now Mass spec vendors have created software for such a determination. There is no good substitute for the MFs. MFs provide a very useful parameter for assessing the assay performance over the long term and specially when we are analyzing unknown samples.	Other methods may be acceptable on exceptional basis, as long as this is scientifically justified.
241-249	19	This approach (based on determining sample concentrations) represents the actual sample analysis more adequately. It should not be limited only to special cases (e.g. on-line sample preparation).	The first approach is in line with the Crystal City paper.
241 - 249	20	Comment: In practice, the commonly used criteria of CV ≤ 20% for MF is not always sensitive enough to identify significant ME. However, I have observed several methods where spiked samples from 1 or 2 of the matrix lots were < 15% different from the mean for other lots and the %CV across all lots was well within acceptance limits. I know of > 5 of these assays that met ME criteria during validation but failed during sample analysis due to ME **Proposed change (if any):** Recommend adding an additional requirement that analyte concentrations in 5	The guideline indicates as criteria a CV of <15%. Regarding the **Proposed change:** Line 241-249 is not mentioning matrix factor. Included is a criteria for CV and every lot.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		of the 6 individual lots evaluated must fall within ±15% of nominal and that additional matrix effect evaluations (such as post-column infusion and/or phospholipid transition experiments) should be performed if these criteria are not met	
241-249	45	For determining the concentrations, one would require calibration standards to be prepared in one of the lots. Which lot should one consider as the calibration backbone. Concentration determination suggested here is equivalent to determination of precision and accuracy of low QC in 6 different lots. Instead of determining the concentrations, the statistics determined for the peak areas of analyte, IS and their ratios would provide a better assessment of the MF and IS normalized MF.	It is considered that this is independent of the back-bone matrix, as variability is requested between lots. The back-bone is the constant factor.
242, 247, 250	11	Comment: does "batch" mean "lot", i.e. 6 lots of matrix? If yes, it is better to use lot for consistency and to avoid confusion.	The text has been revised and states now 'lots'.
247-249	41	Comment: If one batch shows a deviation of >20%, more investigation should be done. This is irrelevant. A single sample exceeding the criteria does not have any statistical value (e.g. compare with a single QC sample failing the criteria in bioanalysis). **Proposed change (if any):** The evaluation should be done over the mean of the six results.	This is agreed. The text has been revised accordingly.
252	7	Comment: If a formulation for injection to be administered to the subjects or animals contains excipients known to be responsible for matrix effects, for instance polyethylene glycol or polysorbate, the potential impact of the excipient on the analytical method should be taken into consideration. This would allow more flexibility in case the concentration of the excipient is very low and/or a stable-labeled internal standard is used. In addition there is no need to use blank matrix from animals treated with the excipient. A matrix effect may	It is agreed that the matrix may be spiked in case there is no metabolism. This is indicated in the guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		also be detected by spiking the excipient to a normal blank matrix or a QC sample.	
252	23	Matrix effect with the excipient: as the excipient is known during the drug development, this will be performed as an additional part of the validation?	This can be done by a partial validation before the study.
252-254	37	Comment: This point is not clear. The list of excipients, which have to be evaluated, should be attached to the guideline. In case of comparative studies and both products containing the same excipients in similar quantity this matrix effect experiments should not be necessary and this should be clearly stated. The minimum quantity of excipient should be defined, for which evaluation is necessary (e.g. daily dose).	A fixed excipients list is not applicable for a guideleine, as this list is subject to changes/updates. This implies that the concentration of the excipient/metabolites is the same. Furthermore, if there is an excipient effect, the obtained concentrations are incorrect. This is not acceptable.
252-256	19	After an appropriate sample preparation the interference of excipients with the method is very rare. This study should be limited to very special cases. Additionally, availability of this kind of matrices is very problematic. For animal studies the excipients can vary in a wide range, it would entail lots of investigations/revalidations.	Excipient interaction should be demonstrated to know if the results obtained are correct.
252 - 256	20	Comment: As a general policy, additional evaluation of ME during MD has proven to be extremely beneficial. Techniques such as post-column infusion (with infused analyte concentrations at or near the LLOQ) and monitoring of the most common phospholipid transitions during MD can effectively guide chromatographic development to ensure resolution of analyte and IS from strongly enhancing or suppressing matrix components. The routine use of these types of techniques in MD will not totally eliminate the possibility of ME but will significantly decrease the likelihood of significant ME in the final method **Proposed change (if any):** Recommend the use of techniques to monitor ME including, but not limited to, post-column infusion and phospholipid transition monitoring experiments as a routine MD tool during	Method development is not the scope of the guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		bioanalytical LC-MS/MS MD	
252-256	21	Matrix from animals that have been dosed with blank vehicle could be obtained during the conduct of a nonclinical study and a separate study should not be required. **Proposed change (if any):** We recommend that this be omitted or that matrix from control animals be used.	This is acceptable. The guideline states only that it should be evaluated/demonstrated. The method to do so may differ.
252-256	22	Comment: Point 4.1.8 Matrix Effect After an appropriate sample preparation the interference of excipients with the method is usually very rare. We would recommend to specifically limit this type of study on a case by case basis. It is also worth noting that the availability of such matrices can prove problematic. E.g. for animal studies, the excipients can vary in a wide range and a systematic implementation of such studies would entail lots of investigations and re-validation work.	Excipient interaction should be demonstrated to know if the results obtained are correct. That is the basic request. By omitting such evaluations, the accuracy of the obtained concentrations is unclear. Mistakes in concentrations obtained during preclinical studies may have a great impact on decisions (for instance on safe toxicity concentrations in humans).
252-256	23	This requirement should be limited to clinical studies. The type of excipient used during the pre-clinical phase is very variable and each change would require a complement to the validation.	These excipients are typical for preclinical studies. Furthermore, mistakes in concentrations obtained during preclinical studies may have a great impact on decisions (for instance on safe toxicity concentrations in humans).
252 - 256	24	Comment: There are ethical considerations to be taken into account to obtain matrix of this nature from animal and even human. **Proposed change (if any):** Please consider re-phrasing: If a formulation for injection to be administered to the subjects or animals contains excipients known to be responsible for matrix effects, for instance polyethylene glycol or polysorbate, the potential impact of the	The proposal is considered not acceptable. Matrix effect may have a large effect on the outcome of a study and therefore this should be evaluated.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		excipient on the analytical method should be taken into consideration.	
252-256	36	The issue of matrix effects of formulation excipients is a complicated one and needs to be considered on a case by case basis. It should be largely left to scientific judgement as to how to deal with these issues based on the specific circumstances. Also, it is not always possible to get matrix from subjects/animals dosed with just the excipient prior to a study. Therefore this requirement should be removed. **Proposed Change:** Delete "The matrix used for this evaluation……or transformed in-vivo.	Excipient interaction should be demonstrated to know if the results obtained are correct. That is the basic request. By omitting such evaluations, the accuracy of the obtained concentrations is unclear. Mistakes in concentrations obtained during preclinical studies may have a great impact on decisions (for instance on safe toxicity concentrations in humans).
252-256	41	Comment: It is usually impossible to test these items in validation, as often the exact formulation is not yet known (i.e. still to be tested in a clinical study). Furthermore, how should this blank matrix be obtained? Should additional studies be executed with just blank formulations? This is ethically unacceptable. Furthermore, typically those samples are owned by pharmaceutical companies, thereby almost eliminating the possibilities of a CRO to do these assessments. **Proposed change (if any):** If this is really needed, it could be done on a study-to-study base, using the placebo samples (if placebo formulation is exactly the same as for active).	Excipient interaction should be demonstrated to know if the results obtained are correct. That is the basic request. By omitting such evaluations, the accuracy of the obtained concentrations is unclear. Mistakes in concentrations obtained during preclinical studies may have a great impact on decision (for instance safe toxicity concentrations in humans). Furthermore, normally blank matrix is obtained in studies.
252-256	48	Comment: - In case of the formulation contains excipients known to be responsible for matrix effect: The investigation of excipients interference is it compulsory in each validation study even if the blank matrix is a precious one (rare; ethic consideration) - In absence of any known matrix effect, this interference study should it be systematically performed in validation study?	Excipient interaction should be demonstrated to know if the results obtained are correct. That is the basic request. By omitting such evaluations, the accuracy of the obtained concentrations is unclear. Mistakes in concentrations obtained during preclinical studies may have a great impact on decisions (for instance safe toxicity concentrations in humans).

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		The absence of the investigation of this parameter would it lead to a not compliant file and so conducted the competent authorities to request complementary experiments?	
254-256	37	Comment: The list of excipients, which have to be evaluated or the list of excipients, which do not have to be evaluated, should be attached to the guideline.	A list is considered not applicable for a guideline, as this list is subject to regular updates and changes
254-256	45	This might be true for small animals and i.v. administration, but not for p.o. and large animals and humans and sophisticated methods (no testing for impact of excipients is needed in the latter cases)	Comment noted.
254-256	47	Comment: Regarding obtaining matrix from subjects that contain only the excipient. Obtaining matrix from human subjects that contains only the excipient(s) may be difficult (for instance, if the study does not have a placebo arm where the subjects are dosed only with the formulation without the drug). There may be objections from IRBs with dosing subjects with excipients solely for the purpose of obtaining matrix. **Proposed:** Recommend that this section be reworded in light of the Comments provided.	Excipient interaction should be demonstrated to know if the results obtained are correct. That is the basic request. By omitting such evaluations, the accuracy of the obtained concentrations is unclear.
257	40	Comment: In section 4.1.9, a list of stability is provided. However, "large batch injection", which is a very critical measure for LC/MS method, is not included. The LC/MS instrument is very sensitive, at the same time not always stable, if the parameter is not well set. Validation batches are generally a lot smaller than batches for sample analysis. To ensure the instrument parameter is well set for the sample analysis, a batch representative to a similar size from the sample batch (~100) should be injected during the validation stage. **Proposed change (if any):** Add, after line 285:	It is considered that this is covered, as the guideline recommends accuracy and precision evaluation of a full batch size and evaluation of autosampler stability.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		For LC/MS method, a batch representative to a similar size of the study sample batch should be injected to evaluate and ensure the parameters are robust for the entire batch. The batch can consist of QC samples from any inter-day precision. Peak area ratio variation evaluation involves comparing low, mid and high QC samples injected at the beginning of a batch with the same samples re-injected in the middle and at the end of the batch. Acceptance Criteria: At least 67% of QC samples at each injection location (beginning, middle and end) must have % R.E. within ±15.0%.	
257 - 261	24	Comment: Stability investigation of IS in matrix and in solution is not needed because instability will become evident during method development. LT Stability data should be available when the study data are reported rather than at the start of the study. We would like EMA to think about the first section because it is contradictory in itself: (initial versus nominal concentrations) We would like EMA to specify further what is meant by stability could not be proved by literature data. We would like EMA to think about to assign blood stability to method development and to allow to compare here to actual concentration rather to nominal since the outcome will direct the assay development and for nominal concentration measurement not method is available (plasma or serum method only). The 15% deviation criteria represents the limit of the chromatography based methods only, we would like EMA to include the LBA method criterion as well and to think about to allow a more scientific approach to cover for "slightly" unstable metabolites and other more statistically based approaches of the assessment of stability. **Proposed change (if any):** Wording: Reads now: "… the conditions applied to the stability tests, such as sample matrix, materials storage and analytical	Stability of IS is needed. These data can be obtained during development. The text is revised, indicating that long term stability data should be available before the study report is issued. For consistency, 'initial' has been revised into 'nominal'. Reference to literature is not acceptable and data should be provided by carrying out stability studies in the laboratory. The text has been revised for clarity. The method to evaluate this is not in the scope of this guideline. It is recognised that chromatographic methods and ligand binding assays are differing to such an extent that combining recommendations and criteria would be confusing. Therefore a separate section has been introduced on ligand binding assays. It is considered that the method in the guideline is clear. A compound is not stable within a certain time period when the criteria are not met. This cannot be changed by the statistical method. The text has been revised and includes now container

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		conditions …" Should read: "… the conditions applied to the stability tests, such as sample matrix, container materials and storage and analytical conditions …"	materials.
257-299	2	Comment: • Consider fixing the number of freeze-thaw cycles per default to 3 (as FDA). If the freeze-thaw cycles need to be adapted to the actual number of cycles in the study then retrospective validation may be needed. • Incurred samples stability should be encouraged when possible. Incurred samples stability represents real conditions. Consider having long-term stability conducted based on spiked QC samples, and when possible or necessary using incurred samples. Acceptance criteria should be based on ISR criteria. • Clarification is required regarding the need to demonstrate internal standard stability in the matrix. The frequent use of stable label internal standards for mass spectrometry would obviate the need for retesting the same chemical moiety a second time. • When using stable label internal standards that are the same salt or free base form and in the same solvents as the drug reference material, the stability tests for the drug would also provide coverage for the internal standard and eliminate the need for retesting the internal standard. • We request clarification regarding the need to demonstrate internal standard stability in the matrix. Internal standards are not stored with samples and therefore stability determination is not required. IS should in general provide a	-It is considered that the guidance is clear. There is always the possibility to cover this issue by partial validation. Furthermore, the FDA guideline is also subject to changes at the moment. The text has been revised, indicating that study samples may be used in addition. This is correct. The text has been revised accordingly. This may be acceptable if justified. This is correct. The text has been revised accordingly. Container drug interaction may be different and can only be noticed at low concentrations. However, this is not a stability

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		reasonable uniform response throughout the run, but that too is under debate (see 'Validation topics with no consensus' in Crystal city 3 whitepaper, AAPS Jr. 2007: 9(1), article 4).	issue.
		• Stock and working solution stability should only be tested when the solvent system changes during dilution of the stock to working solution. A chemical would not be expected to change stability at different concentrations.	This is always an issue whether the sample is prepared correctly. The same applies for the freshly spiked and prepared sample. A separate section has been introduced on ligand binding assays.
		• How is the freshly prepared calibration curve qualified to ensure that it was correctly prepared before it is used to assess the stability samples? Criteria are for small molecule only – no specific guidance is given for large molecule.	This criterion is taken into account to ensure that the QCs are prepared correctly.
		• Suggest removing requirement to analyze low and high stability QCs immediately after preparation since the comparison would be made to the nominal concentration. Although analysis immediately after preparation may give useful information on the accuracy of QC sample preparation, it should not be a requirement.	The text has been revised and indicates that: 'For small molecules it is considered acceptable to apply a bracketing approach, i.e. in case stability has been proved for instance at -70°C and -20°C, it is not necessary to investigate the stability at temperatures in between. For large molecules (such as peptides and proteins) stability should be studied at each temperature at which study samples will be stored.'
		• To avoid discussion around detailed interpretation of different storage conditions it should be made clear that stability should be investigated for storage conditions that bracket all that may apply to study samples.	The text is revised, indicating that long term stability data should be available before the study report is issued.
		• The recommendation that evaluation of long term stability is carried out before the start of the actual study would impact the ability to start studies and the amount of long-term stability needed is unknown, whereas aligning with the FDA requirement of having to demonstrate stability for at least the duration the study	For consistency, 'initial' has been revised into 'nominal'. This is considered not acceptable. 48 h means that the samples have been stored and as such cannot be used for comparison. Freshly spiked samples should be used for this purpose.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		samples were stored should be sufficient. • Line 261 discusses 'Any deviation from the initial concentration that does occur must be within acceptable limits', while the sentence related to QC stability defines that stability is measured against the nominal concentration. Needs clarification. **Proposed change (if any):** • Replace the "freshly prepared" with more specific statement such as "standards prepared from reference standards within 48 hours and stored at the most conservative condition …." for small molecule. • Line 265-267 change to: Stability of the analyte in the studied matrix is evaluated using at least triplicate samples of the low and high QC samples which are analysed after the applied storage conditions that are to be evaluated. • Line 294-295: Delete the sentence 'It is recommended that evaluation of long term stability is carried out before the start of the actual study'.	The text is revised only with regard of deleting 'IS'. The text is revised, indicating that long term stability data should be available before the study report is issued.
257-299	13	Comment: The concurrent evaluation of the stability of reference standards or test samples during the testing of clinical samples, which is proposed by the guideline, seems to be incompatible with the design of a clinical study.	The comment is unclear.
257-299	15	For small molecules: If shown that the analytes are stable at e.g. -20 °C, lower temperatures should be accepted without further testing. Incurred sample stability (ISS) should be investigated with reference to the first analytical result, since ISS might give hints on the presence of unstable metabolites.	The text has been revised and indicates that: 'For small molecules it is considered acceptable to apply a bracketing approach, i.e. in case stability has been proved for instance at -70°C and -20°C, it is not necessary to investigate the stability at temperatures in between. For large molecules (such as peptides and proteins) stability should be studied at each temperature at which study samples will be stored.'

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
257-299	26	Comment: Line 267: Why is it important to evaluate the analyte concentration in the QC immediately after its preparation if stability is being evaluated relative to the nominal concentration? This would be more relevant if stability were evaluated relative to initial result. Initial analysis would, however, serve to confirm correct preparation of the QCs. Line 269: An alternative approach might be to verify that the initial freezing and storage of a plasma spikes overnight does not affect its concentration if assayed the next day. Line 275: Probably not necessary if the same solvent is used for the stock and working solution preparation and stock and working solutions are stored under the same conditions. Line 296: This process needs to be defined clearly unless results are to be used "for exploratory use only". For example, do you need a separate validated assay for spikes in blood that are used as part of a plasma validation? Care needs to be taken to define the "reference" concentrations in this type of experiment. Presumably the intent is to use a concentration determined after application of optimal processing conditions (time, temperature, etc.) compared to concentrations determined after one or more set of stressed conditions. With blood processing to plasma, spiked (nominal) concentrations in blood cannot be used as the reference concentration due to red blood cell/plasma distribution. **Proposed change (if any)**: Clarify proposals	This criterion is taken into account to ensure that the QCs are prepared correctly. It seems that using freshly prepared samples will be avoided here. This is considered not acceptable. If considered not necessary by a CRO, a test can be omitted, if supported. The guideline does not indicate a request for systemic analysis. Demonstration of stability in blood may be needed on a case by case situation depending on the compound. The text has been revised, indicating that demonstration of stability may be needed on a case by case basis.
257-299	36	**4.1.9 Stability** There appears to be some contradictions and lack of clarity in the terms employed in this section (initial versus nominal concentrations). Greater clarity and definitions of the terms and requirements is sought from	For consistency, 'initial' has been revised into 'nominal'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		the agency for this section.	
257-299	47	Comment: Regarding the assessment of stability. This section does not include guidance that when there is more than one analyte in a sample, whether it be two different analytes or an analyte and its metabolite(s), that stability of each analyte in the presence of the other(s) needs to be considered in the context of the study design. Additionally, this section does not include procedures to consider in regard to anticoagulants and anions/cations. **Proposed:** Recommend that this section be reworded in light of the comments provided.	The text has been revised into: 'In case of a multi-analyte study and specific for bioequivalence studies, attention should be paid to stability of the analytes in the matrix containing all the analytes.'. The text has been revised and includes now anticoagulants.
257-299	52	Comments: The concurrent evaluation of the stability of reference standards or test samples during the testing of clinical samples, which is proposed by the guideline, seems to be incompatible with the design of a clinical study.	The comment is not understood.
258-299	22	Comment: Point 4.1.9 Stability All stabilities should be calculated regarding the nominal value instead of comparing them with time 0. In the section: "The QC samples are analysed... the obtained concentrations are compared to the nominal. The deviation should be within ±15%" (in lines 267-269), clarification is needed as to whether this is deviation compared to the mean or individual value?	For consistency, 'initial' has been revised into 'nominal'. Furthermore, the text has been revised regarding the deviation: the guideline states now 'mean concentration'.
257	23	The stability of the blood after the collection and the separation of the plasma/serum should not be considered part of this method validation but part of a	If considered not necessary by a CRO, a test can be omitted, if supported.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		more global method development.	
257		Should a criteria of stability be taken into account for the analysis of the "time zero" with respect to the nominal value? Please could you add some precision.	This criterion is taken into account to ensure that the QCs are prepared correctly. The nominal concentration should be within 15% (in line with QC accuracy criteria).
258-269		Concerning the deviation with time, should this be determined with respect to the initial value or with respect to the nominal value?	For consistency, 'initial' has been revised into 'nominal'.
269		Concerning the deviation for the stabilities, should this be determined with respect to each of the initial values or with respect to the mean value calculated at each level of concentration?	The text has been revised clearly indicating now 'nominal concentration'.
257-299		Section 4.1.9 Stability; there are some contradictions in this section (initial versus nominal concentrations: line 260 versus line 269) and unclear.	For consistency, 'initial' has been revised into 'nominal'.
		In addition, we don't understand why would we need to study the stability of the IS in the matrix. We propose that blood stability should be evaluated versus reference, instead of versus nominal.	The text has been revised and 'IS' has been deleted.
		For assessment of stability of the analyte and IS in the studies matrix, are you expecting the results of QC samples to be compared to nominal concentrations, or should the results for the 'after storage' QC samples also be compared to the results for the freshly prepared QC samples?	The text has been revised and clearly indicates now that this should be done using the nominal concentrations.
		Stability cannot be proven by literature data. Agreed! Is it possible to refer to stability studies obtained by another method or obtained at another site by the same method (e.g. at a CRO).	This may be acceptable in case this applies to the same site using another method but applying the same storage conditions including same container and volume.
		Proposed Change: Replace " Any deviation from the initial concentration that does occur must be within acceptable limits" by "Any deviation from the nominal	For consistency and clarity, 'initial' has been revised into

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		concentration that does occur must be within the nominal limits".	'nominal'.
260	7	Comment: Initial concentration should be replaced by nominal concentration **Proposed change (if any):** Initial concentration should be replaced by nominal concentration	For consistency and clarity, 'initial' has been revised into 'nominal'.
260	21	Initial concentration should be replaced by nominal concentration **Proposed change (if any):** initial concentration should be replaced by nominal concentration	For consistency and clarity, 'initial' has been revised into 'nominal'.
260	47	Comment: With regard to the statement that during stability analysis any deviation from initial concentration that does occur must be within acceptable limits, we are unclear as to what the intent of this statement is. **Proposed:** Recommend that further clarification be included in the guidance.	For consistency and clarity, 'initial' has been revised into 'nominal'. Furthermore, this sentence only indicates that after sampling handling and/or storage, the concentration should be comparble to the nominal concentration, within a certain limit.
260-261	37	Comment: We suggest the use of confidence intervals for stability evaluation [U. Timm, M. Wall, D. Dell, J. Pharm. Sci. 74 (1985) 972-977; P. Rudzki, A. Leś, Acta Pol. Pharm. 65 (2008), 743-747]. This statistical approach clearly increases quality of data. Acceptance criteria 85-115%. **Proposed change (if any):** Any deviation from the initial concentration that does occur must be within acceptable limits. It is suggested to use 90% confidence intervals for statistical evaluation of stability data, but other statistical approaches may be used if justified.	It is considered not very usefull to apply confidence intervals for such a low number of observations. Furthermore, the applied method described in the guideline is a generally accepted method. The proposed change is therefore not agreed.
261 and 267	16	Comment: Line 261 discusses 'Any deviation from the	For consistency and clarity, 'initial' has been revised into

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		initial concentration that does occur must be within acceptable limits'; while the sentence is related to QC stability, it defines that stability is measured against the nominal conc. **Proposed change (if any):** Please clarify line 261 and keep the assessment as measured against nominal concentration to be consistent with existing practices.	'nominal'.
262-264	41	Comment: Analytical conditions should be similar to those used in actual study samples. Does this imply that e.g. the collection of blank matrix used in the validation should be done the same as clinical study samples will be drawn? Sampling conditions are often not known prior to validation. **Proposed change (if any):** Revision of text.	The guideline refers to stability tests and the applied storage conditions which should cover the conditions during sampling handling.
264	8	Comment: The statement implies that at least one set of stability QC samples undergo all the stability stressors at once (e.g. FT, ST and LT) and not only each stability evaluation separately. Is this required?	The guideline refers to stability tests and the applied storage conditions which should cover the conditions during sampling handling. This does not imply that stability should be proven in one test covering start and end of sampling handling and storage.
264	14	Comment: For co-medications in drug-drug interaction studies, especially well known "old" drugs (i.e. warfarin) literature data on stability in sample matrix should be allowed to be used as reference. **Proposed change:** To delete the sentence.	This is not agreed. It should be ensured that the conditions in house are covered; this cannot be done by literature data as the specific conditions in the test facility are not known.
264	19	"Stability cannot be proven by literature data." In-house stability results of stock and working solutions should be allowed to refer to when the method is validated for another matrix/species and the same solutions (with equal concentrations and solvents) are used. Stability of stock and working solutions should not be re-tested.	If considered not necessary by a CRO, a test can be omitted, if supported. If stability of stock and working solutions have been evaluated under the same conditions, a retest is considered not necessary.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
264	22	Comment: Point 4.1.9 Stability "Stability cannot be proven by literature data." Experience shows that in-house stability results of stock and working solutions can be referred to when the bioanalytical method is validated for an other matrix/species and the same solutions (with equal concentrations and solvents) are used. Stability of stock and working solutions should therefore not have be re-tested. We would recommend this situation to be depicted in the final text.	If considered not necessary by a CRO, a test can be omitted, if supported. If stability of stock and working solutions have been evaluated under the same conditions, a retest is considered not necessary.
264	45	Does "literature data" mean cross-reference to open scientific literature or does it include also internal company reports?	Literature data refers to open scientific literature. Reference to internal company reports is acceptable, as long as the conditions applied are the same.
265	7	Comment: There is no need to assess stability of IS in matrix **Proposed change (if any):** Please delete "..and IS..."	The text has been revised and 'IS' has been deleted.
265	8	Comment: Stability of the IS in matrix is required, however, as stated previously, IS is not spiked into the calibration curve. It is added as part of the extraction process, and the precision and accuracy data will determine if there is an IS extraction issue.	The text has been revised and 'IS' has been deleted.
265	21	No need to assess stability of IS in matrix, **Proposed change (if any):** please delete "..and IS..."	The text has been revised and 'IS' has been deleted.
265	32	Comment: similar to Brussels comments ie the stability of the IS in the study matrix need not be evaluated as it is spiked fresh on the day of analysis and any decomposition will be observed in the blank matrix spiked only with the IS.	The text has been revised and 'IS' has been deleted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
265	36	**Comment :** Internal standard (IS) stability is requested to be evaluated. This is not clear as to why this needs to be done and is not common best practice unless specifically if required to be demonstrated during sample proceeding and instrument analysis.	The text has been revised and 'IS' has been deleted.
265	45	Stability determinations for the IS are not needed (see comment to lines 92-94)	The text has been revised and 'IS' has been deleted.
265	47	Comment: Regarding the stability of the internal standard in matrix. When determining analyte stability in matrix, concentration comparisons are made to freshly prepared samples, however there are no such comparisons for the internal standard. Would the comparison expectation be an internal standard peak area of a stored sample to that of a freshly prepared sample? Additionally, what would be the acceptance criteria for demonstration of stability? **Proposed:** Recommend describing how internal standard stability in matrix will be determined immediately after preparation and after the applied storage conditions.	The text has been revised and 'IS' has been deleted.
265-267	14	Comment: Investigations of the stability of the internal standard in the studied matrices are not required since the internal standard will be spiked freshly to every sample during sample clean-up. Decomposition of the internal standard would be observed since also blank matrix samples only spiked with the internal standard will be evaluated within every run (e.g. system suitability test). Furthermore, for isotope-labelled internal standards the physical-chemical properties are the same as for the non- labelled drug. **Proposed change:** Please delete the requirements on stability investigations for the internal standard.	The text has been revised and 'IS' has been deleted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
265-267	15	Investigations of the stability of the internal standard in the studied matrices are not required since the internal standard will be spiked freshly to every sample during sample clean-up. Decomposition of the internal standard would be observed since also blank matrix samples only spiked with the internal standard will be evaluated within every run (e.g. system suitability test). Furthermore, for isotope-labelled internal standards the physical-chemical properties are the same as for the non-labelled drug. **Proposed change:** Please delete the internal standard from the requirements on stability investigations.	The text has been revised and 'IS' has been deleted.
265-267	18	Comment: IS are added to the sample prior to processing therefore stability tests of IS in the matrix are relevant only for in-process stability and room-temperature stability prior to processing.	The text has been revised and 'IS' has been deleted.
265-267	21	Suggest removing requirement to analyze low and high stability QCs immediately after preparation since the comparison would be made to the nominal concentration. Although analysis immediately after preparation may give useful information on the accuracy of QC sample preparation, it should not be a requirement. **Proposed change (if any):** Stability of the analyte ~~and IS~~ in the studied matrix is evaluated using at least triplicate~~s~~ samples of the low and high QC samples which are analysed ~~immediately after preparation and~~ after the applied storage conditions that are to be evaluated. Other approaches, if based on statistically-sound bases, are also acceptable.	This criterion is taken into account to ensure that the QCs are prepared correctly. The nominal concentration should be within 15% (in line with QC accuracy criteria).
265-267	39	It should be emphasized that IS is normally added to the samples during sample processing. Thus, it has no sense to test its stability in the matrix.	The text has been revised and 'IS' has been deleted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): Delete the words "and IS".	
265-269	5	Comment: If standard curves show e.g. a systematic but within the acceptance criteria lying "trend" for the QC-H or QC-L to a positive or negative deviation from the nominal value then also within a stability investigation the investigated QC-H or QC-L will show this "trend". But these observations/results are not related to drug stability. Therefore, a comparison of the concentration of freshly prepared QCs with the obtained concentration of stored QCs for stability evaluation would eliminate measurement related deviations.	As indicated, the nominal concentration should be used.
265-267	6	Comment: IS are added to the sample prior to processing therefore stability tests of IS in the matrix are relevant only for in-process stability and room-temperature stability prior to processing.	The text has been revised and 'IS' has been deleted.
265 – 269	35	Comment: For stability evaluation the nominal concentration is not clearly defined. Is it the back calculated concentration of QC samples which are analysed immediately or is it the theoretical value **Proposed change (if any):** The definition for 'Nominal' could be included in the 'Definition' section or in the text at the relevant sections	The nominal concentration should be used, which is the theoretical value.
265-269	45	Other approaches exist in the scientific literature (see for instance Journal Pharmaceutical Sciences 74 (1985) 972) which have more sound statistical basis. Please add sentence allowing different approaches	The applied method described in the guideline is a generally acceppted method and considered most simple and sufficient.
265-271	21	Use the term "stability samples" instead of QC samples. Levels are OK. Suggestion to rename to stability samples instead of QC samples. QC samples are for batch acceptance and for evaluation of accuracy and precision. Stability samples are for stability assessment.	The term QC sample is considered clear.
266	36	**Comment**	It is difficult to include a time period in the guideline as this

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Clarity is sought on what the agencies definition of immediately is in the requirement to "QC samples which are analysed immediately after preparation"	may be questioned. For sure this is not 48 h after preparation, and also not 12 h after preparation.
266	41	Comment: QC samples analysed immediately after preparation and after storage. Why should this de done? Concentrations are calculated against nominal values after storage. **Proposed change (if any):** delete requirement of analysis directly after preparation.	This criterion is taken into account to ensure that the QCs are prepared correctly. The nominal concentration should be within 15% (in line with QC accuracy criteria).
266	45	**Proposed change (if any):** Rename samples from "QC" samples to "stability samples".	The term QC sample is considered clear.
266-269	32	Comment: similar to EBF presentation, the stability samples should be named as such ie stability samples and not QC samples as this causes confusion. QC samples are used to accept analytical batches and to evaluate precision and accuracy.	The term QC sample is considered clear.
267-268	38	Comment: Does "Freshly prepared calibration standards" means Freshly spiked or just freshly preparation and injection.	The text has been revised and indicates now , freshly spiked'.
267-269	37	Comment: In case of no stability issues there is no need to use freshly prepare calibration standards. In case of serious stability problems there is no chance to pass acceptance criteria of stability by using unstable calibration standards. So there is no use to state in guideline that freshly prepare calibration standards must be used, it leads to no improvement in measurement results quality. It seems to be an overregulation. The comparison of stability samples to nominal concentration may be not most reliable stability estimate, because actual concentration of QC samples used for stability assessment may differ from nominal concentration. Therefore, in our opinion, stability experiments should be based on comparison of two data sets: 1) reference samples (e.g. freshly prepared samples)	Both methods have there pro's and con's. The method in the guideline is considered most accurate.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		2) stability samples (after storage for desired time in suitable conditions). **Proposed change (if any):** The QC samples are analysed against a calibration curve, ~~obtained from freshly prepared calibration standards,~~ and the obtained concentrations are compared to the concentrations of reference samples (e.g. freshly prepared) or nominal concentrations.	
268	36	Use of term QC samples can be misleading. **Proposed Change:** Substitute phrase "QC samples" for "stability samples"	The term QC sample is considered clear.
269	19	What do you mean on "The deviation should be within ±15%."? It should be clarified. **Proposed change (if any):** We suggest the following acceptance criteria (the same as for accuracy and precision test): the mean accuracy should be within ±15% of the nominal value, the CV value should not exceed 15% and the ±15% criteria should also be fulfilled for at least 50% of the individual QC samples tested per concentration level.	The text has been revised for clarity. It is considered that estimation of the mean accuracy is sufficient for the evaluation of stability. However this does not mean that borderline results cannot be questioned.
269	21	Add acceptance criteria for ELISA (± 20%)	A separate section has been introduced on ligand binding assays.
269	23	The stability is evaluated using 3 replicates. The text states "The deviation should be within ± 15%" Is the deviation of 15% applicable to the individual or mean value? The deviation should be evaluated in terms of precision and accuracy. The CV value should not exceed 15% for the QC stability. For accuracy if the value should be within 15% of the nominal value, any grade of non-stability is assumed for analytes. **Proposed change (if any):** The CV value should not exceed 15% for the QC samples (precision). The mean	The text has been revised for clarity. It is considered that estimation of the mean accuracy to be within 15% is sufficient for the evaluation of stability. However this does not mean that borderline results cannot be questioned.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		accuracy value should be within 20% the nominal values for the QC samples.	
269	45	Add acceptance criteria for ELISA (± 20%)	A separate section has been introduced on ligand binding assays.
269	47	Comment: Regarding the use of "deviation." **Proposed:** Recommend consistent use of terminology here vs. lines 174, 175, 202, 203, 207, and 208, for example.	The text has been revised for clarity.
269	51	Comments: for stability replace deviation with %Bias or %RE, also add to section that you are measuring individual QCs not the group and that 1) at least 2 of the three samples must be within the bias tolerance and 2) that the group %CV must be less than or equal to the tolerance **Proposed change (if any):**Replace Deviation with %Bias And clarify group vs individual values	The text has been revised for clarity. It is considered that estimation of the mean accuracy to be within 15% is sufficient for the evaluation of stability. However this does not mean that borderline results cannot be questioned.
269	51	Comments: There is no description of CV for group of stability samples. **Proposed change (if any):** Add description of group CV	It is considered that estimation of the mean accuracy to be within 15% is sufficient for the evaluation of stability. However this does not mean that borderline results cannot be questioned.
270	23	What are the acceptance criteria for the stability of the stock and working solutions?	The mean concentration at each level should be within ±15% of the nominal concentration.
270	23	Regarding working solution stability test, what concentration levels should be tested? Should this stability test done if working solutions are prepared freshly everyday of use?	The text has been revised to clarify this issue.
270-271	19	There are no acceptance criteria for stock and working solution stability. Please include.	The mean concentration at each level should be within ±15% of the nominal concentration.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): We suggest to be within reference ±10%.	
270 - 273	24	Comment: To avoid discussion around detailed interpretation of different storage conditions it should be made clear that stability should be investigated for storage conditions that bracket all that may apply to study samples. **Proposed change (if any):** Stability studies should investigate different storage conditions that are at or bracket all storage conditions over time periods that equal or exceed those applied to the actual study samples.	The text has been revised to clarify this issue.
271	23	Stability should be tested only if necessary since calibration curve, may be obtained from freshly prepared stock and working solutions.	Comment noted.
272,273	17	Comment: It is unclear if this statement includes planned retains, which could lead to very long studies. **Proposed change:** Stability studies should investigate different storage conditions over time periods that equal or exceed those applied to the actual study samples except planned retains.	It is not clear what is meant by planned retains.
272-273	21	It is too restrictive to determine and scientifically not needed to determine the storage stability at the same storage temperature as the study samples. Suggest revising it to include stability at or above the storage temperature of the study samples, unless it is known that colder temperatures interfere with the stability To avoid discussion around detailed interpretation of different storage conditions it should be made clear that stability should be investigated for storage conditions that bracket all that may apply to study samples. **Proposed change (if any):** Stability studies should	The text has been revised to clarify this issue. This is considered not applicable to LBA.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		investigate different storage conditions *that are at or bracket all storage conditions* over time periods that equal or exceed those applied to the actual study samples.	
274 - 285	24	Comment: We would like EMA to specify what they are meaning with stock and working solution stability and what acceptance criteria should be applied to them. We would as well ask EMA to define the difference of their meaning of bench top stability of processed samples	The text has been revised for clarity. The mean concentration at each level should be within ±15% of the nominal concentration.
275	16	Comment: Stock and working solution stability should only be tested when the solvent system changes during dilution of the stock to working solution. A chemical would not be expected to change stability at different concentrations. When using stable label internal standards that are the same salt or free base form and in the same solvents as the drug reference material, the stability tests for the drug would also provide coverage for the internal standard and eliminate the need for retesting the internal standard. **Proposed change (if any):** Please provide specific cases where stability testing is not needed based on sound science.	The text has been revised to clarify this issue.
275	20	Comment: IS stock and working solution stability is generally not performed because 1) it is not recommended in the current Guidance documents, and 2) because IS is added to all samples (rationale: if interferences are present, all samples should be similarly affected). Unfortunately, ME are not always consistent from sample-to-sample. I am aware of several instances where IS stocks were used for extended periods (up to 1 year or more) with no stability evaluation. If any unknown degradation products were present in these older solutions, the impact of these components on assay accuracy for actual samples would	It is agreed that this should also apply to IS. The text has been revised to include IS.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		most likely remain unknown. It is good scientific practice to include IS stability **Proposed change (if any):** Recommend including requirements for IS stability in stock and working solutions	
275	45	If stock and working solutions are in the same solvent, abandon working solution stability investigation	The text has been revised to clarify this issue.
275	47	Comment: Regarding the types of stability tests evaluated. **Proposed:** Although not a stability assessment per se, stock weighing does not appear to be included in the guidance. The relevance to the acceptable use of a similar stock for the preparation of standards and QC samples within a matrix is not described. Furthermore, since some laboratories justify the averaging of stock weighings and resulting concentrations, recommend that the practice of stock weighing, preparation, storage stability, and use be clarified.	This is considered out of the scope of this guideline. The text has been revised. The QC samples should be spiked independently from the calibration standards, using separately prepared stock solutions, unless the nominal concentration(s) of the stock solutions have been established.
275-285	37	Comment: to facilitate understanding, the stability test should be listed in the following order: firstly - required, then – if applicable.	The text has been revised according to the proposal.
276-277, 279	8	Comment: Although these evaluations are examples, some analytes cannot be left at room temperature, and all sample processing is done at a colder temperature e.g. on an ice/water (4°C nominal). Statements are very specific.	This is considered specific situations, which cannot be covered by this guideline.
278-279	23	What is the difference between these two sentences? What does "bench top" really mean particularly with respect to duration?	This mean stability testing of the same sample, but at different storage temperatures. Furthermore, the text has been revised for clarity.
279	37	Comment: widely used term "short-term stability" should be used (as approved by FDA). **Proposed change (if any):** ~~bench top~~ short-term	The text has been revised and indicates now 'short-term'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		stability of the analyte in matrix at room temperature	
280-281	45	It is too restrictive to determine and scientifically not needed to determine the storage stability at the same storage temperature as the study samples. Suggest revising it to include stability at or above the storage temperature of the study samples, unless it is known that colder temperatures interfere with the stability	The text has been revised and indicates that: 'For small molecules it is considered acceptable to apply a bracketing approach, i.e. in case stability has been proved for instance at -70°C and -20°C, it is not necessary to investigate the stability at temperatures in between. For large molecules (such as peptides and proteins) stability should be studied at each temperature at which study samples will be stored.'
282-283	45	Clarify here the term bench top stability of the processed sample	The text has been revised for clarity.
284	11	Comment: add "or unprocessed" after "processed" because it may be unprocessed samples for on-line extraction	The text has been revised for clarity. Furthermore, this is considered covered by short term stability of the analyte in matrix.
284	36	**Comment** Text related to on-instrument and auto sampler stability are not necessarily relevant to all analytical platforms (e.g. LBA methods) **Proposed change** Text should include phrase "as applicable to the assay methodology"	The text has been revised for clarity. This is considered not applicable to LBA.
286-288	50	Comment: Regarding freeze and thaw stability, the QC samples are stored and frozen in the freezer at the intended temperature and thereafter thawed at room temperature. After thawing, samples are re-frozen again applying the same conditions. Here, thawing duration for freeze and thaw cycle is missing. For freeze and thaw stability evaluation, it is appropriate to incorporate the "Thawing duration for each freeze and thaw cycle" in this section. Here thawing duration indicates the duration for which samples kept on bench at room temperature.	The text has been revised and indicates now 'after complete thawing'.
286-290	45	12 hours for frozen samples in freeze-thaw cycle stability investigation: Too detailed, allow to facilitate three cycles a day but one has to make sure that	This alleged contradiction is not confirmed.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		samples are fully frozen and thawed in each cycle	
286-290	47	Comment: Regarding the freeze and thaw stability, the freeze/thaw practice is not complete. Specifically, the thawing period is not defined. Since the initial freezing period of stability QC samples is not defined, a laboratory has carte blanche to freeze QC samples for any period of time, even less than 12 hours. Although it has become the industry practice to freeze long-term stability samples at both -20ºC and -70ºC, there remains no guidance as to doing so. Note that FDA CDER has voiced its opinion regarding this, and similarly recommends assessing long-term stability at both temperatures. **Proposed** Recommend that an acceptable period of thawing be described. At minimum, method validation needs to define a priori the thawing period, and the period should reflect a sufficient period of time that a sample requires for complete thawing. Furthermore, the conditions upon which a typical incurred sample is thawed should be reproduced during method validation. Note that there have also been discussions on simulating the thawing period of samples in the laboratory prior to sample preparation as a best practice. Recommend that the initial freezing period be documented within the guidance to be at least 24 hours. Recommend long-term stability be assessed at both -20ºC and -70ºC.	The text has been revised and indicates now 'after complete thawing'. The guideline states at each cycle, meaning that samples should have been frozen 12 h before thawing. The text has been revised and indicates that: 'For small molecules it is considered acceptable to apply a bracketing approach, i.e. in case stability has been proved for instance at -70°C and -20°C, it is not necessary to investigate the stability at temperatures in between. For large molecules (such as peptides and proteins) stability should be studied at each temperature at which study samples will be stored.' The text has been revised and indicates now 'after complete thawing'. No further guidance is considered needed.
288	41	Comment: first cycle 12 hours frozen contradict the FDA Guidance (24 hours frozen) **Proposed change (if any):** Harmonize with FDA	This alleged contradiction is not confirmed.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Guidance.	
288-289	14	Comment: What is the scientific rational behind the 12 hours frozen time? Would a shorter time interval for SMOL not also be appropriate?	This is a regulatory set criteria.
288- 289	23	Freeze and thaw stability The samples should be completely thawed before they are refrozen **Proposed change (if any):** At each cycle, when completely thawed, samples should be frozen for at least 12 hours	The text has been revised and indicates now 'after complete thawing'.
289-290	31	Comment: The guideline states that the number of cycles in the freeze-thaw stability should equal or exceed that of the freeze/thaw cycles of study samples. **Proposed change:** It is suggested to specify a standardized number of cycles, e.g. like FDA Guidance Bioanalytical Method Validation (2001) at least three cycles.	It is considered not necessary to include a standardised number, as during study sample analysis this number may be also exceeded.
291-295	14	Comment: In many cases it is not possible to evaluate long term stability before start of a study, since the duration of a study is very often unknown. Alternatively, it should be acceptable to use study samples for the evaluation of long term stability. In this case the first valid concentation value of the study sample can be taken as the starting value for long term stability investigations. **Proposed change:** A more appropriate wording to take long term stability investigation into consideration, also from practical point of view, would be: "*Regarding long term stability of the analyte in matrix stored in the freezer: The QC samples should be stored in the freezer under the same storage*	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued. The use of study samples for evaluation of stability is considered not acceptable, as the nominal concentration in this sample is not known. As such it cannot be ensured that the value at day 0 is correct. The proposed change is not agreed.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		conditions and at least for the same duration as the study samples. The investigation should be initiated before start of the actual study. For the evaluation of the long term stability the use of study samples may also be acceptable if a valid concentration value for the study sample is available which can be used as starting point for the stability investigation. A prerequisite for this approach is that for the value used as starting point stability for this truncated time interval is guaranteed and stability data are available."	
291 - 295	24	Comment: We think that stability investigations from study samples (ISS) can deliver very useful information about stability and integrity of the analyte, and sometimes (endogenous compounds) this is the only way to assess and to guide method development. To our understanding it covers the impact of labile drugs and/or metabolites and freeze-thaw cycles as well. Acceptance criteria as ISR could be applied in case of stability investigations with study samples.	The use of study samples for evaluation of stability is considered not acceptable, as the nominal concentration in this sample is not known. As such it cannot be ensured that the value at day 0 is correct. However, additional tests can be carried out using study samples.
291 – 295	35	Comment: For long-term stability since there are no QC samples, which are analysed immediately, it is necessary to define the criteria for comparison of the results obtained for long-term stability. **Proposed change (if any):** For long term stability cconcentrations of all the stability samples could be compared to the mean of back-calculated values for the standards at the appropriate concentrations from the first day of long-term stability testing	The guideline indicates that the QC sample concentration. should be evaluated at the start of stability testing (day 0).
291-295	45	1. If -20° stability is ok, this covers -70°C for small molecules. 2. Long term stability data should be available when study results are reported or can be still under investigation but this must be clearly stated in the report. 3. Study samples can give valuable information about stability of the drug and associated metabolites which can not be obtained from	1. The text has been revised and indicates that: 'For small molecules it is considered acceptable to apply a bracketing approach, i.e. in case stability has been proved for instance at -70°C and -20°C, it is not necessary to investigate the stability at temperatures in between. For large molecules (such as peptides and proteins) stability should be studied at each temperature at which study samples will be stored.'This

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		spiked matrix 4. It is not always practical that the long term stability can be always known before the start of the actual study. Some flexibility is required here. Stability experiments can be performed concurrent or shortly after the study samples analysis	is considered not applicable to LBA. 2. The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued. 3. The use of study samples for evaluation of stability is considered not acceptable, as the nominal concentration in this sample is not known. As such it cannot be ensured that the value at day 0 is correct. However, additional tests can be carried out using study samples. 4. The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
293	23	"not acceptable to use study samples for evaluation of long term stability": We still believe this should be kept as an option when appropriate (e.g. known presence of a metabolite that can degrade back to the analyte but that is not available as reference to conduct stability studies); in that case, comparison would be based on initial concentration instead of nominal which is unknown.	The use of study samples for evaluation of stability is considered not acceptable, as the nominal concentration in this sample is not known. As such it cannot be ensured that the value at day 0 is correct. However, additional tests can be carried out using study samples.
293-295	21	We disagree with the following proposal "For the evaluation of the long term stability it is not acceptable to use study samples, as the nominal concentration is unknown, and can therefore not be used as reference". We consider that the evaluation of stability in actual study samples and not only in QCs should be kept as an option when appropriate (e;g. known presence of a metabolite that can degrade back to the analyte but that is not available as reference to conduct stability studies.	The use of study samples for evaluation of stability is considered not acceptable, as the nominal concentration in this sample is not known. As such it cannot be ensured that the value at day 0 is correct. However, additional tests can be carried out using study samples.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Even if we do not know the nominal concentration, the %Diff with reference gives information regarding stability. **Proposed change (if any):** Replace "For the evaluation of the long term stability it is not acceptable to use study samples, as the nominal concentration is unknown, and can therefore not be used as reference." by "Alternatively long term stability could be carried out using actual samples."	
294	23	Evaluation of long term stab before study start would be nice but will be unrealistic in term of planning in most situations. What matters in the end is that the study results are covered with appropriate stability data and that the stability required to cover the period of the study is available before the finalization of the study report. The sentence lines 294-295 could be skipped or modified in this sense.	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
294	36	**Comments** The guidance recommends evaluating long term stability before the start of the actual study This may be impractical and not predictable. Furthermore it may require use of resource and generate unnecessary data that is not relevant to the requirements of the project **Proposed change** Storage stability data of sufficient duration must be available to confirm that all reportable sample concentrations in a study were determined within that validated storage duration**.**	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
294-295	19	Evaluation of long-term stability before start of the actual study (trial sample analysis) should not be required. **Proposed change (if any):** This sentence should be deleted or revised as the stability results must be available before finalisation of the actual study report at	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		the latest.	
294-295	21	The recommendation that evaluation of long term stability is carried out before the start of the actual study would impact the ability to start studies and the amount of long-term stability needed is unknown, whereas aligning with the FDA requirement of having to demonstrate stability for at least the duration the study samples were stored should be sufficient. **Proposed change (if any):** Delete the sentence 'It is recommended that evaluation of long term stability is carried out before the start of the actual study'.	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
294-295	29	Comment: Due to practical and logistical reasons, it is often difficult to establish long term stability of analytes in matrix, before the start of the actual study. Hence this recommendation may be removed. Even the AAPS article (Bansal S, DeStefano A. Key Elements of Bioanalytical Method Validation for Small Molecules. *AAPS Journal*. 2007; 9(1): E109-E114), cites that the long term stability in matrix can be done post validation and the data can submitted at a later date. **Proposed change (if any):** It is recommended, although not mandatory to perform the evaluation of long term stability in matrix, before the start of the actual study.	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
294-295	31	Comment: It is recommended that evaluation of long term stability should be carried out before the start of the actual study. Long term stability results may be not sufficient for the duration of the actual study. Hence extension of long term stability should start as soon as the duration of study is known. **Proposed change:** Long term stability should be evaluated before start of the actual study and in case of need to extent long term stability, results should be available before analysing	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		study samples.	
294-295	32	Comment: It is not always feasible to complete long term stability studies before the start of sample analysis, therefore it is proposed to add "where feasible" after the word "recommended".	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
294-295	34	Comment: The duration of long term stability could be one or two years and the data obtained during this period are often appended to the initial validation report. **Proposed change (if any):** It is recommended that evaluation of long term stability covers the duration of storage between sampling and sample analysis.	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
294-295	37	Comment: The recommendation is definitely wise and of general knowledge, but it is not always possible to determine long-term stability before the start of the study. As post-study determination of long-term stability does not effect quality of data, we propose to delete this recommendation. **Proposed change (if any):** ~~It is recommended that evaluation of long term stability is carried out before the start of the actual study.~~	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
294-295 371-372	6	Comment: Why should the evaluation of long term stability be carried out before the start of the actual study? And in general, why should the entire validation be performed before the analysis of study samples? The validation results are valid regardless whether they are obtain before or after analysis of study samples; criteria for evaluation are predefined and cannot be changed based on obtained results. It is risky for analytical laboratory to perform some validation experiments during or after analysis of study samples, but if these validation experiments have satisfactory results they should be accepted.	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): Do not require to perform all validation experiments particularly long-term stability before beginning of analysis of study samples.	
294-295 371-372	39	Why should the evaluation of long term stability be carried out before the start of the actual study? And in general, why should the entire validation be performed before the analysis of study samples? The validation results are valid regardless whether they are obtain before or after analysis of study samples; criteria for evaluation are predefined and cannot be changed based on obtained results. It is risky for analytical laboratory to perform some validation experiments during or after analysis of study samples, but if these validation experiments have satisfactory results they should be accepted. **Proposed change (if any):** Do not require to perform all validation experiments including long-term stability before beginning of analysis of study samples.	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
295	7	Comment: "It is recommended that evaluation of the long term stability is carried out before the start of the study." This is not practical in all circumstances. Although it is acknowledged that primarily spiked stability samples should be used to assess analyte stability study samples may provide very helpful "add-on" information, i.e. with respect to metabolites present. There, it should be acceptable to provide data on life sample stability if felt necessary. **Proposed change (if any):** Please rephrase to "It is recommended that evaluation of long term stability is initiated before the start of the actual study because stability data must be available	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		upon reporting of study data.”	
295	23	According to the draft guideline, long-term stability tests are recommended to be carried out **BEFORE the** start of the actual study? To avoid differences in approach, shouldn't it be defined in the guideline what periods of storage are considered to be classed as 'long-term' and should be performed before start of the study? Also we feel that this requirement be modified to identify that the “long term stability be demonstrated for the period required and be available before the finalization of the report containing the results of the study.	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
295	49	It is not possible to perform long-term stability before the start of the study. Long-term stability is part of within-study validation, not pre-study validation. **Proposed change (if any):** Long-term stability should be performed as soon as possible after completion of analytical phase. This allows for accurate calculation of time required to be evaluated.	The text has been revised and the guideline indicates now that long term stability data should be available before the study report is issued.
296	36	Clarification to the statement is required. The text “Sufficient attention should be paid to the stability of the analyte in the sampled matrix directly after blood sampling of subjects and further preparation before storage, to ensure that the obtained concentrations by the analytical method reflect the concentrations of the analyte in the subject at the moment of sampling” is unclear since it starts with "matrix stability" but then refers to blood sampling.. **Proposed change** Suggest simplifying the statement to “the stability of the analyte in the collection matrix (e.g. blood) throughout the sampling and processing procedures should be demonstrated “.	The text has been revised and indicates: ‘A demonstration of this stability may be needed on a case-by-case basis, depending on the structure of the analyte.’.
296-297	23	Could you clarify “Sufficient attention should be paid to the stability of the analyte in the sampled matrix directly after blood sampling of patients and further preparation	The text has been revised and indicates: ‘A demonstration of this stability may be needed on a case-

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		before storage". Is it the evaluation of the stability in blood during sample preparation, like centrifugation? Regarding the stability between the sampling, the processing and the storage is there any proposal on ways of doing it?	by-case basis, depending on the structure of the analyte.'.
296-297	47	Comment: Regarding the statement that attention needs to be paid to the stability of the analyte in the sampled matrix directly after blood sampling of subjects and further preparation before storage, to ensure that the obtained concentrations by the analytical method reflect the concentrations of the analyte in the subject at the moment of sampling. It is felt that insufficient guidance is given here. Additionally, the guidance only discusses blood sampling and does not offer guidance on other types of sampling such as urine and tissue. **Proposed:** Recommend that consideration of sample collection stability be addressed more thoroughly to also include other matrices such as urine and tissue. Stability at collection is not necessarily part of the bioanalytical method process, therefore it should be documented if the stability of sample collection was not determined, and the potential impact on the method should be documented. This is particularly relevant to urine. Note that if the collection of the sample is not controlled, then the stability of the matrix at freezing cannot be assumed to be controlled, and thus validated. In the cases of where urine and solid tissue are not assessed for stability at collection, then the bioanalytical method cannot be fully validated. In actuality, the method can only then be qualified to the extent of that as defined in the method report.	The text has been revised and indicates: 'A demonstration of this stability may be needed on a case-by-case basis, depending on the structure of the analyte.'. Regarding other matrices, like urine and tissues: these are normally used as the final matrix, which is mostly not the case for blood.
296-299	19	The need for this test should be evaluated on a case by case basis (e.g. it might be appropriate for esters). Additionally, it would be useful if this section was more	The is considered not appropriate for a guideline, but for a SOP.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		detailed. The guideline should suggest a procedure how to do this test. Acceptance criteria might be different from those of other stability tests, for instance blood stability might be evaluated against reference, instead of against nominal concentration.	
296-299	21	This is a whole blood stability requirement. Failed stability can be complicated to rectify since many factors are involved. Whole blood contains a mixture of plasma and red blood cells, validation experiments measure stability of the compound in plasma. The compound is found in the plasma; therefore, it would be stable in the plasma component of whole blood. While understanding the stability of the analyte in the sampled matrix directly after blood sampling is important in theory, it is difficult to control how samples are handled "in the field". Because this process is very difficult to control, we feel that there is little value in generating additional data to understand its effect **Proposed change (if any):** We recommend removing this requirement.	The text has been revised and indicates: 'A demonstration of this stability may be needed on a case-by-case basis, depending on the structure of the analyte.'.
296-299	22	Comment: Point 4.1.9 Stability Stability of the analyte during sample collection (in whole blood) is of importance. We would recommend including in the final text more details about the methodology to be applied and also that blood is considered of acceptable quality for 48 hours after blood collection when kept at 4ºC.	The is considered not appropriate for a guideline, but for a SOP. Furthermore, the text has been revised and indicates: 'A demonstration of this stability may be needed on a case-by-case basis, depending on the structure of the analyte.'.
296-299	45	1. It is unclear how this can be achieved; e.g. blood is processed to serum (sample matrix); are we being asked to determine stability in whole blood for the duration of the blood-to-serum process? 2. While understanding the stability of the analyte	1. If needed, the stability of the analyte in blood should be evaluated. 2. The text has been revised and indicates: 'A demonstration of this stability may be needed on a case-by-case basis, depending on the structure of the analyte.'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		in the sampled matrix directly after blood sampling is important in theory, it is difficult to control how samples are handled "in the field". Because this process is very difficult to control, we feel that there is little value in generating additional data to understand its effect.	
296-299	49	Analytical laboratories have no influence/control over samples directly after sampling. This is the responsibility of the clinic.	The responsibility of the accurateness and reliability of the results is the responsibility of the sponsor and/or the applicant.
296-299	50	Blood Stability: additional clarity should be provided to avoid any ambiguity, it is suggested to prove stability of analyte in whole blood for 2-6 hours based on time taken to prepare plasma from blood especially for drugs which may show binding to or may penetrate to blood cells. This may ensure an extended and long term performance of a bioanalytical method.	If needed, the stability of the analyte in blood should be evaluated. However, the text has been revised and indicates: 'A demonstration of this stability may be needed on a case-by-case basis, depending on the structure of the analyte.'.
299	11	Comment: remove the extra "." at the end.	Text has been revised.
299	50	Comment: Moment of sampling.. (Redundancy in full stop). Superfluous dot (.) should be removed	Text has been revised.
300	23	It would be useful if this section was more precise and identify the different types of validation for pre-clinical and clinical situations, for example: Typical bio-analytical method changes that fall into this category include, but are not limited to: · Bio-analytical method transfers between laboratories or analysts · Change in analytical methodology (e.g., change in detection systems) · Change in anticoagulant in harvesting biological fluid · Change in matrix within species (e.g., human plasma to human urine) · Change in sample processing procedures	Some examples have been added. However it will not cover all situations.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		· Change in species within matrix (e.g., rat plasma to mouse plasma) · Change in relevant concentration range · Changes in instruments and/or software platforms · Limited sample volume (e.g., pediatric study) · Rare matrices · Selectivity demonstration of an analyte in the presence of concomitant medication · Selectivity demonstration of an analyte in the presence of specific metabolites.	
300-307	2	Comment: • Partial validation into another matrix/species should be added. It's time and resource consuming if a full validation is needed e.g., for all species. • Partial validation not needed if new calibration range is within the originally validated range, i.e., higher LLOQ and/or lower ULOQ.	Some examples have been added. However it will not cover all situations. This is agreed, as long as it does not affect the regression model
300 - 307	24	Comment: We think, there should be clearer definitions whether a full or a partial validation is needed (e.g., method applied to other species, applied to a same matrix but collected on other anticoagulant, applied on limited sample volume, chnage of the calibration range, ...) . Modifications or updates of bioanalytical method requiring a partial validation should be clearly pre-defined in a SOP. The way those parameters will be evaluated as well as their acceptance criteria should also be described.	Some examples have been added. However it will not cover all situations. It is agreed that this should be described in a SOP.
300-307	26	General: Some indication of what constitutes an acceptable partial validation should be indicated. Line 304: For transfer of the method to another laboratory, a full validation in the new laboratory should be required.	Transfer to another laboratory is considered subject to a partial validation, however, as indicated, it depends on the nature of the applied changes.
300-307	28	Comment: other cases exists, in our opinion, when a partial validation could be used when the different	Some examples have been added. However it will not cover all situations.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		species/same matrice are studied: case of non-separative techniques involving the assay of a total metal like ICP-MS, optical ICP. So the guideline should leave some flexibility on the application of a partial validation, on a case-by-case basis and providing that the sponsor will justify it's strategy in the submission (IMPD, MA dossier, etc)	The guideline indicates this already, by stating 'depending on the nature of the applied changes'.
302	4	Comment: delete 'to' in the word 'onto' in that line	The text has been revised.
303	23	**Proposed Change:** Complete the possibility to perform partial validations when there is a change of species, change of anti-coagulant, selectivity with metabolites or concomitant medication, change of extraction method… (as described in the report of the conference "Bio-analytical Method Validation – a revisit with a decade of progress" Workshop held in Arlington, VA, 2000).	Some examples have been added. However it will not cover all situations.
305	47	Comment: Regarding the use of "justified." **Proposed:** Recommend that "rationalized and tested" be used rather than "justified."	It is considered that the term 'justified' is justified.
306	23	How many Accuracy and Precision runs are required?	The required number of runs depends on the nature of the applied changes.
306-307	37	Comment: The re-confirmation of response function (calibration curve) seems to be the most important. We propose to limit accuracy-precision experiments to within-run investigation. **Proposed change (if any):** In most cases, provision of calibration curve, within-run accuracy and precision data or relevant additional stability data on the modified issue may be sufficient.	The required supportive data from a partial validation depends on the nature of the applied changes.
308 Cross validation	36	**Comments** Clarification is required as to what constitutes "data obtained from different study sites" This suggests that analysis conducted in a multi-centre study requires cross	The text has been revised and indicates now 'different laboratories'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		validation experiments. This is contradictory to the wording given in line 312 which expects this to be done ahead of study sample analysis. The proposed differences for acceptance criteria are too stringent at 15% **Proposed change** Reword section to define cross validation being required when two separate assays are used to report sample concentration data from within the same study samples Acceptance criteria should either be based on a statistical tool for determining "true" differences or should be set to be no greater than 30% difference**.**	The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.'.
308-316	2	Comment: • Is this criterion indicated for both large and small molecules or would large molecule criterion be different? • Cross validation should be indicated only when data obtained from different methods are compared within and across studies and when data are obtained within a study from different sites applying the same method. Cross validation should be performed on spiked sample matrix and study samples applying criteria similar to those used for incurred sample reanalysis. **Proposed change (if any):** • Replace first sentence in lines 309-310 by: "Where data are obtained from different methods within and across studies or when data are obtained within a study from different sites applying the same method, comparison of those data is needed, and a cross validation of the applied analytical methods should be carried out." • Replace last sentence in lines 315-316 by: "For at least 2/3 of the samples the difference	A separate section has been introduced on ligand binding assays. The text has been revised in accordance with the proposed change. The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		between the two measurements should not exceed 20% (30% for LBA)."	whether they can be compared and used.'.
308-316 315	23	How many batches/ validation days are recommended for cross or partial validations? In which cases do we need to consider "cross validation"? Does this apply only to studies for which analyses are performed in two different laboratories or are there other cases? Cross validation: For cross validation we propose to add standards and unknown samples to QC samples and that a statistical analysis should be performed. **Proposed Change:** Modify to "difference between two measurements should not exceed 40% ".This is more in line with the other method validation criteria (i.e. accuracy and precision within 15%, if one method accuracy 115%, and the other 85%, then difference is 30%, not 15%). However, we think that the acceptance criteria appear too stringent and should be in line with the other method validation criteria. Why not apply the acceptance level of plus or minus 20% as in the re-analysis criteria?	For partial validation, the required number of runs depends on the nature of the applied changes. For cross validation: one run is sufficient, in line with ISR. It is considered that for cross validation, the use of QC alone is acceptable. Additional samples may be added. The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.'.
308 - 316	24	Comment: This is a new definition of cross-validation as opposed to the definition by the FDA. Actually we would prefer to follow FDA guidance requirements. The requirement of "The difference between the two measurements should not exceed 15%" is far more stringent than "The difference between measured and nominal QC concentration should not exceed ±15%". Please specify LBA criteria: ±20% compared to nominal concentration?	Line 309-311 have been revised for clarity. The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Based on incurred samples, the use of 20% (30%) acceptance criteria which is in agreement with ISR criteria, two/third within 20%(LC-MS) and 30%(LBA) to original value. In addition, the FDA guideline does not specify that the cross validation assessment should be done before the analysis of study samples. Is cross-validation only applicable when different methods are used within the same study? If not, please specify for which studies and when cross-validation of methods is required (e.g., different techniques used in the two methods) Cross-validation strategy and acceptance criteria should be pre-defined in a SOP and/or in study protocol.	A separate section has been introduced on ligand binding assays. The guideline indicates that cross validation should be done in advance of study sample analysis, if possible. Comment noted.
308-316	26	Line 309 (and Line 311): Change "study" to "laboratory" to reduce possible confusion with clinical (or nonclinical) study conduct sites. Line 313: At a minimum, the same set of QC samples should be analysed by both laboratories. Pooled incurred samples should also be included in the cross-validation, if available. Line 315. When several QCs and pooled incurred samples are evaluated, there should be some allowance for a limited number of results to be greater than 15% apart, just as for QCs in an analytical run, where 2/3 must pass. **Proposed change (if any):** Modify as above. Include formula for calculation of the closeness of agreement of the two different results	Line 309-311 have been revised for clarity. It is considered that for cross validation, the use of QC alone is acceptable. Additional samples may be added. The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.'
308-316	41	Comment: Cross-validation done with QC samples only. **Proposed change (if any):** It should be possible to do cross-validation with incurred samples.	It is considered that for cross validation, the use of QC alone is acceptable. Additional samples may be added.
308-316	45	1. Full validation will ensure the assay performance at each of the analytical site, and will allow	1. Comment noted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		comparison of the literature data. 2. This is a new definition of cross-validation as opposed to the definition by the FDA. Actually we would prefer to follow FDA guidance requirements. Use of 20%(30%) acceptance criteria which is in agreement with ISR criteria, two/thirds within 20%(LC-MS) and 30%(ELISA) to original value 3. It should be specified what percentage of the samples should be within acceptance.	2. Line 309-311 have been revised for clarity. 2/3. The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.'. A separate section has been introduced on ligand binding assays.
309-310	14	**Proposed change:** Propose to rephrase to: "*Where data is obtained from multiple bioanalytical study sites/or employing multiple analytical methods within one bioanalytical site for a single study, a comparison of those data is required, and a cross validation of the applied analytical methods should be carried out."*	The proposed change is considered to complex. Line 309-311 have been revised for clarity.
309-310	15	Please make a clear distinction between cross- and transfer-validation. **Proposed change:** It should not read „Where data are obtained from different study sites…." but "Where data are obtained by different analytical methods…" Because different study sites (=laboratories) are addressed in 4.2 Partial validation and it refers to "method transfer" and the respective partial validation requirements.	For transfer validation a partial validation is needed. This is also indicated under section 4.2 Partal validation.
309-316	21	Cross validation is indicated only when data obtained from different methods are compared within and across	Line 309-311 have been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		studies and when data are obtained within a study from different sites applying the same method. Cross validation should be performed on spiked sample matrix and study samples applying criteria similar to those used for incurred sample reanalysis. **Proposed change (if any):** replace line 309-316 by: Where data are obtained from different methods within and across studies or when data are obtained within a study from different sites applying the same method, comparison of those data is needed, and a cross validation of the applied analytical methods should be carried out. Differences in sample preparation or the use of another analytical method may result in different outcomes between the study sites. Cross validation should be performed in advance of study samples being analysed if possible. For the cross validation, the same set of QC samples and study samples should be analysed by both analytical methods/sites. If it is not feasible to analyze the same set of QC samples and/or study samples (e.g. due to a stability issue), an alternative approach should be employed to compare the methods. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used. For at least 2/3 of the samples the difference between the two measurements should not exceed 20% (30% for LBA).	The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.' A separate section has been introduced on ligand binding assays.
309-316	23	This requirement should be limited to clinical studies.	Cross validation is not limited to clinical studies, but also to toxicokinetics.
312-313	47	Comment: Regarding the statement that "For the cross-validation, the same set of QC samples should be analyzed by both analytical methods." We feel that this should be a recommendation rather than a requirement. This approach works well where the cross-validation activities take place in the same facility, but what if the cross-	The text has been revised and indicates now: 'For the cross validation, the same set of QC samples or study samples should be analysed by both analytical methods.'

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		validation activities are taking place at several sites? According to the EMEA guidance requirements, QC samples would have to be made by one facility and shipped to other facilities. This may introduce un-anticipated bias due to numerous stability considerations **Proposed:** Recommend that the guidance be worded as "Criteria for the acceptance of the cross-validation activities should be determined and justified a priori." This would allow individual laboratories to utilize and implement other methods of determining cross-validation acceptability.	
315	9	**Proposed change (if any):** If stated, more detail should be given on the specification for the cross-validation exercise. In addition, due to the inherent greater variability of ligand binding assays, different (wider) criteria need to be stated for such assays. Suggest to also indicate that statistical approaches may also be used for comparison on data.	Line 309-311 have been revised for clarity. The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.' A separate section has been introduced on ligand binding assays.
315-316	7	Comment: We find the acceptance criteria too stringent and suggest to modify to "***difference between two measurements should not exceed 30%".*** This is more in line with the other method validation criteria (i.e. accuracy and precision within 15%, incurred sample reproducibility allows 20% deviation of mean **Proposed change (if any):** Please change to "***difference between two measurements should not exceed 30%".***	The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.' A separate section has been introduced on ligand binding

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			assays.
315-316	14	**Proposed change:** Propose to rephrase to: "*Deviation of the back calculated QC concentrations should not exceed 15% of the nominal value. If appropriate, a cross validation with study samples can be an alternative approach, especially to take possible interferences with metabolites into account. The difference between the two measurements of study samples should not exceed 15 % of the mean value.*"	The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.'
315-316	15	The acceptance criteria are not appropriate. **Proposed change:** It should be stated that 2/3 should not exceed a difference of more than 15 % or 20 % from the mean (similar to the ISR criteria?).	The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.'
315-316	37	Comment: The acceptance criteria are definitely too strict. There is virtually no chance that each pair will pass this criteria. There are significant differences in equipment, personnel and reference material (i.e. calibration samples) between sites. How many QC samples and which QC concentration levels should be analysed? We recommend to apply similar acceptance criteria as for incurred samples. **Proposed change (if any):** The difference between the two individual values obtained in different study sites should be within 20% of their mean (30% for ligand-binding assays) for at least 67% of the measurements.	The following criteria have been included: 'For QC samples, the obtained mean accuracy by the different methods should be within 15% and may be wider, if justified. For study samples, the difference between the two values obtained should be within 20% of the mean for at least 67% of the repeats. The outcome of the cross validation is critical in determining whether the obtained data are reliable and whether they can be compared and used.'
316	11	Comment: How about cross validation at the LLOQ level? In this case, it should be 20%.	It is considered that the same calibration curve range is applied. Accuracy data for the LLOQ will be obtained during run validation of the analysis for the QC samples. As indicated before, the acceptance criteria for LLOQ is within 20% of the

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			nominal concentration.
317 - 320	24	Comment: We prefer a major revision of this section to provide more clarity and to align with state of the art LBA validation practice, which is very much based upon the Crystal City III workshop report and the papers written by DeSilva et al. For LBAs, the inter-batch variance component is usually a greater contributor to the overall variability than the intra-batch variance component. It is recommended that at least 2 independent determinations be made for each validation sample per assay run across a minimum of 6 independent assays runs (balance validation design).	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
317-364	2	Comment: • It is unclear which parts of sections 4.1- 4.3 apply to LBA and which are LC-MS specific. • What acceptance criteria would be applied for specificity? An indication of what recovery of LLOQ spikes is acceptable should be given. As described, this experiment would be considered "selectivity" for large molecules. Specificity is typically tested by spiking in a similar therapeutic molecule and demonstrating that the method measures only the therapeutic of interest. • Line 325-326: This recommendation seems to be too extensive. Please acknowledge that in many cases it is not feasible to obtain different batches of an analyte (especially in early (pre)clinical stages) or to obtain the analyte in different phosphorylation states and glycoforms to study the effect of phosphorylation and glycosylation in the assays. It is unclear what is	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		meant with the metabolic differences in phosphorylation. • Lines 324-328: Elaborated how the batch-to-batch differences of proteins should be evaluated, i.e. can be covered later by partial validation with a comparison between the two batches (precision-accuracy runs with two batches). • Lines 331-340: It is unclear under what circumstances there would be pre-existing antibodies to the analyte at sufficient concentration and affinity to impact on an assay. For some analytes there may be endogenous binding proteins, soluble receptors and other molecules that could bind to the analyte, if this is what is meant then it would be helpful to redraft this section. • Line 331: Where plasma is the analytical matrix include one lot each haemolyzed and hyperlipideamic plasma in the total of 10 lots. • Lines 350-354: more details could be provided on calculating the total error. And when to use it. • Total Error should not be a LBA requirement if it is not a requirement for LCMS. • Lines 358-360: "Evaluation should not only address chemical and physical properties, but also biological integrity (i.e., maintenance of antibody binding affinity)" – this is vague in terms of what antibody is referred to and what may be the practical implementation of this requirement since affinity determination in biological matrix is clearly outside of typical scope of bioanalytical method validation. This	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		statement is most appropriate for small molecules. Binding differences can be evaluated for large molecules, but it is not clear how an LBA method can evaluate chemical and physical properties of biopharmaceutical drugs. • For LBAs qualification of standard calibrators and QCs would be required if subsequent batches demonstrate differences in glycosylation and metabolic differences in phosphorylation. Some differences in glycosylation are expected as the biologic drug is heterogeneous. When the therapeutic is produced in a different cell line/host then it is relevant to discuss differences. Many protein biotherapeutics are produced in cells and thus, just like in the body, have a range of acceptable and active forms. • For LBAs specificity and selectivity would be demonstrated if the recovery is within 20% of the spiked concentration, except for spikes at the LLOQ and ULOQ concentration which should not exceed 25%. • For the validation of between-run accuracy and precision, at least two determinations, independently prepared, per concentration per run at LLOQ, low, medium, high QC and ULOQ samples from six runs analysed on at least two different days by different analysts should be evaluated. The between-run accuracy and precision should be within ±20%%, except for LLOQ and ULOQ which should not exceed 25%.	
317-364	26	Line 320: Change "antibody" to "binding reagent". The term "ligand-binding assay" is specifically chosen because all binding reagents in these assays are not antibodies, but may include transport or binding proteins, aptamers, etc. Lines 325-326: How is it proposed that validation	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		addresses these differences? Line 328: Change "antibody" to "binding reagent". Line 329: The word "mainly" is vague. Provide clarity. Line 330: Change "antibody" to "binding reagent"; change "structural" to "structurally". Please suggest how this demonstration may be achievable with a protein analyte with undefined biotransformation products (i.e., catabolites). Lines 331-340: these are reasonable objectives which may be practicable to achieve. However, they do not address the statements on Lines 329-330. Line 341: It is not clear how a reference standard should be selected "in such a manner that specificity is ensured and binding characteristics are durable…..". There is no way to ensure selectivity, since reference standards themselves are typically heterogeneous. Certainly a change in reference standard as development proceeds is likely to provide purer material with different binding characteristics. Suggest that the wording be changed to indicate that the purest reference standard available at the time be procured. There also needs to be some mention of the comparability of the reference standard material and that to be administered to animals or humans in the actual study. Line 342: Although awkward, "antibody/antigen complex" should be changed to "binding reagent/analyte complex" (see comment for Line 320) Line 355: This may not always be practical; for instance, in cases where the assay is applied to support of a small nonclinical study or a large clinical study the run sizes may differ.	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Lines 357-359: This suggestion is not easy to implement for macromolecules in the concentrations employed in ligand-binding assays. Binding differences can be evaluated, but differences in chemical and physical properties will be much more difficult to demonstrate in dilute solutions. Line 360: Delete "antibody". The key binding reagent need not be an antibody, but may include other binding proteins, aptamers, etc. (see comment for Line 320).	
317-364	47	Comment: Ligand-binding assays (LBAs) are not adequately defined in this section or throughout the draft guidance. For instance, as to whether the draft guidance is supposed to address immunoassays is not clear, which could also implicate ADA/immunogenicity, receptor assays, biomarker assays, etc., needs to be clarified within this document. **Proposed:** For many in the industry conducting ligand-binding assays (LBAs) and macromolecular analyses, it is believed that immunoassays and LBAs are being treated like a footnote or add-on in numerous guidances and White Papers (which is also demonstrated in this particular guidance). There have been suggestions about removing this section altogether, which SQA BASS would not object to, however that would leave a void within the bioanalytical method validation paradigm associated with GLP and human clinical bioanalysis. The period in which a guidance could otherwise be prepared to address LBA and other immunoassays is not predictable, but likely too long. Therefore, there is an immediate need for a guidance document that would extend beyond a stated understanding that the basic principles of validating bioanalytical methods for small molecule also apply to LBAs except for a variation in the types of validation method parameters that require investigation and documentation, and the allowance for	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		looser acceptance criteria because of the unique heterogeneous and other aspects of the immunoassay (supported with scientific rationale). SQA BASS feels that the EMEA guidance offers the ideal opportunity to provide guidance relevant to LBAs, and even other immunoassays, including ADA and biomarker methods, since having such a guidance would facilitate consistent regulatory approaches to ensuring adequate practices, such as scientific peer review and QAU/agency auditing, while allowing for the creative scientific decisions necessary to facilitate effective drug development and appropriate drug submissions. It is the SQA BASS recommendation that a similar approach and language to immunoassays (at least ligand-binding assays) is needed throughout the EMEA guidance as currently described for LC/MS methods.	
317-364	47	Comment: Regarding acceptance criteria associated with replicates. **Proposed:** There is no discussion in this section of the guidance regarding acceptance criteria associated with replicates, which is necessary to document a priori. Although this does not have relevance to LC/MS methods, it is relevant to ligand-binding assays, where samples are prepared in duplicate or triplicate wells. Acceptance criteria may be applied to individual wells (or not) and/or to back-calculated concentrations based on individual wells or the average of signals/concentrations. Recommend inclusion of acceptance criteria of this type since this practice is currently ad hoc in many laboratories and is often modified post-validation without a partial validation because it is not defined in the current FDA guidance or in Crystal City White Papers.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
317 - 369	3	Quality control samples (QC). LBA QCs should be independently prepared in neat matrix, when possible. Must take into account the frequent existence of	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		endogenous proteins and how to handle their contribution. Suggestion to use the term "drug naïve matrix" rather than blank matrix, since previous dosing of animals even after a long wash out period (blank of drug) could harbor anti-drug antibodies which could impact LBA.	
317 - 369	3	A discussion on Sensitivity should be added. The sensitivity of the method is defined once the LLOQ has been validated for both acceptable precision and accuracy. The sensitivity should be reported as the established LLOQ. Very high concentrations of samples or QCs diluted into the curve to be read do not contribute to the sensitivity estimation. Samples may not be diluted below the value established by acceptable performance of the LLOQ.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
317-369	21	This section 4.4 should be more detailed. Specific paragraph on selectivity, carry-over, LLOQ, calibration curve, accuracy, precision, dilution integrity, matrix effect and stability should be included **Proposed change (if any):** To add Selectivity Selectivity should be proven by using at least 6 sources of the appropriate blank matrix, which are individually analysed and evaluated for interference. Absence of interfering components is accepted where the response is less than 25% of the lower limit of quantitation for the analyte. Calibration curve A minimum of six calibration concentration levels should be used, excluding the blank sample (processed matrix sample without analyte and without IS). A relationship which can simply and adequately describe the response of the instrument with regard to the analyte should be applied. The blank and zero samples	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		might be taken into consideration to calculate the calibration curve parameters. Stability Stability of the analyte and IS in the studied matrix is evaluated using at least triplicates samples of the low and high QC samples which are analysed immediately after preparation and after the applied storage conditions that are to be evaluated. The QC samples are analysed against a calibration curve, obtained from freshly prepared calibration standards, and the obtained concentrations are compared to the nominal concentrations. The deviation should be within ± 20%.	
317-369	23	This section 4.4 should be detailed. Specific paragraphs on selectivity, carry-over, LOQ, calibration curve, accuracy, precision, dilution integrity, matrix effect and stability should be included. **Proposed Change:** To add: Selectivity Selectivity should be proven by using at lest 6 sources of appropriate blank matrix which are individually analysed and evaluated for interference. Absence of interfering compounds is accepted where the response is less than 25% of the LOQ for the analyte. Calibration Curve: A minimum of 6 calibration concentration levels should be used, excluding the blank sample (processed matrix sample without analyte and without IS). A relationship which can simply and adequately describe the response of the instrument with regard to the analyte should be applied. The blank and zero samples might be taken into account to calculate the calibration curve parameters. Stability: Stability of the analyte and IS in the studies matrix is	The text is revised and a complete new and separate section has been introduced on ligand binding assays. .

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		evaluated using at least triplicate samples of low and high QC samples which are analysed immediately after preparation and after the applied storage conditions that are to be evaluated. The QC samples are analysed against a calibration curve obtained from freshly prepared calibration standards and the obtained concentrations are compared to the nominal concentrations. The deviation should be within plus or minus 20%. **Ligand binding assays:** QCs on micro-titre plate based assays. Samples are often analysed as two or three replicates on a plate. FDA requires two samples per QC level resulting in two times three wells. This draft guideline asks for three QC levels. Is it sufficient to include one sample per QC level, if each sample is analysed in two or three replicates on one plate?	
317 -369	36	**4.4 Ligand binding assays** Use of antibodies is only one example how an LBA method can be set up. Other reagents exist (e.g. drug target, biotin-streptavidin and other). Antibody use can be listed as an example of assay set up. **Proposed change** The guidance should be modified to clarify that use of antibodies in Ligand Binding Assays is only one example of assay set up. Other protocols are possible and are routinely used. This affects in particular statements from lines 327 on specificity and line 329 on selectivity.	The text is revised and a complete new and separate section has been introduced on ligand binding assays. .
317-369	45	**Proposed change (if any):** 1. Wording: change "macromolecules" by "biotherapeutics" 2. Wording: change "antibody" to "capture" and	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		"detection molecule"	
318 - 320	3	The term "ligand-binding assay" is specifically chosen because all binding reagents in these assays are not antibodies, but may include transport or binding proteins, aptamers, etc. Please edit "antibody" to "binding reagent".	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
318 - 323	3	These sentences could be more accurate since LBA have always (not just nowadays) been used to quantitate macromolecular drugs. Perhaps a more apt phraseology would be: "Ligand-binding assays (LBA, immunoassays) are the method of choice when quantitation of macromolecules is needed. The validation and study sample analysis principles discussed for small molecule methods can only generally be applied to macromolecule analysis. An understanding of the differences and LBA-specific elements also needs to be addressed to ensure acceptable development and validation of LBAs."	Macromolecules can also be analysed by other methods, so LBA immunoassays are not the method of choice.
318-323	45	To be added to chapter 2 (Scope):	Scope is general and includes therefore also these assays.
324 - 326	3	Batch-to-batch differences in glycosylation and metabolic differences in phosphorylation can be handled as a qualification, and not necessarily as part of validation.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
324 - 328	24	Comment: Batch differences in glycosylation or phosphorylation are accounted for in cell line cross over studies. This work almost never takes place during the initial validation due to the fact that there is only 1 process batch available. We have yet to see any significant difference in assay response when comparing two different cell line batches. However, the batch of the reference standard used for preparation of calibration standards and QC:s is always the same as used for dosing in the clinical or non-clinical study. This should be a requirement.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		We have seen differences in the assay specificity when comparing disease state plasma to stock control plasma. This difference can be significant when analyzing samples from patients with an autoimmune disease. The reason for this interference is more than likely due to heterophilic antibodies or RF factors that can bind to monoclonal antibodies which are mouse/rat derived.	
324-328	45	1. It is suggested that it be indicated that the reference material used for the calibrator curves should reflect the molecule being used in the study and measured in the assay. In addition, this material should be well characterized and stable. The characterization of heterogeneous nature of macromolecules should be tested in the analytical groups vs. in the bioanalytical groups. 2. It is not clear in what practical way the batch-to-batch differences in glycosylation and metabolic differences in phosphorylation should be addressed. Batch-to-batch and/or lot-to-lot differences is almost impossible to evaluate during validation as future lots/batches are usually not available. Are we being asked to partially/fully validate a method when reference standard lots change?	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
324-328	47	Comment: Regarding the discussion of issues specific to ligand-binding assays (LBAs). The paragraph is vague and refers only to an issue associated with LBAs, as opposed to providing guidance on the acceptable approaches to validating LBA methods and associated acceptance criteria. **Proposed:** If this document was written to adequately provide guidance for ligand-binding assays, then this ambiguity would be resolved throughout the document.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Recommend the omission of this paragraph and consideration of other points for inclusion in the document.	
325-326	14	**Proposed change:** Propose to rephrase to: "*As long as there are changes in the production process, validation must cover issues such as batch-to-batch differences in glycosylation and metabolic differences in phosphorylation*."	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
325-326	21	This recommendation seems to be too extensive. Please acknowledge that in many cases it is not feasible to obtain different batches of an analyte (especially in early (pre)clinical stages) or to obtain the analyte in different phosphorylation states and glycoforms to study the effect of phosphorylation and glycosylation in the assays. It is unclear what is meant with the metabolic differences in phosphorylation. We prefer to recommend that the standard to be used is the closest as possible as the one given to preclinical species or human. In case of change of batch, a comparison of batches will be performed. **Proposed change (if any):** Replace Lines 325 - 326 "Therefore...phosphorylation." by: Impact of heterogeneity of the analyte should be considered during method development using standards as close as possible as the one given to preclinical species or man. In case of change of batch, a comparison of batches will be performed."	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
325-328	23	We disagree with the sentence "Therefore validation must cover issues such as batch to batch differences in glycolysation and metabolic differences in phosphorylation". We prefer to recommend that the standard to be used is the closest as possible as the one given to pre-clinical species or man. In case of change of batch, a comparison of batches will be performed.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
327	45	**Proposed change (if any):** Remove specificity and use	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		selectivity instead	
329	47	Comment: Regarding the reference and discussion on specificity. **Proposed:** Recommend that specificity be discussed in Section 4.1, also, and encompass all relevant aspects to ligand-binding assays and LC/MS methods.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
329-330	45	No cross-reactivity with structural related compounds" is vague term. In addition, absence of cross-reactivity of particular antibody is not the only factor providing specificity of the assay (e.g. it can be based on the use of two antibodies so that even without absolute specificity of each one the assay is specific).	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
329,331	9	**Proposed change (if any):** Suggest to use term "Selectivity" not "Specificity"	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
329 - 340	3	**Specificity and Selectivity** - For a LBA the specificity of an antibody refers to the ability of the critical antibody reagent to identify the antigen of interest, i.e., the degree of discrimination between what is positive and what is negative. Selectivity is the ability of an assay to measure the analyte of interest in the presence of other constituents in the sample. The specificity of the reagents needs to be defined before going into validation during reagent development and these reagents are then selected for the intended purpose. Generally, LBAs measure analytes directly in the biological matrix without sample pre-treatment (e.g., extraction), and could be susceptible to nonspecific responses caused by reagent cross reactivity with matrix components. Note that the term "metabolite" is not useful for LBA purposes. While there is some degradation of drug, no back conversion, for instance, is observed. Specificity is tested by adding "related molecules" into drug naïve and spiked matrices. This test is difficult to accomplish given that related molecules are not typically	Comment noted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		generated (or easily generated) by chemists, as they are for small molecule analysts. Some address this issue by adding in possible confounding proteins, disease matrices etc, e.g., RA patient sera, other non-specific abs of same isotype, analyte variants, etc. Selectivity is tested by spiking the therapeutic in a large number of individual samples (greater than 6). Disease state samples should be included, when appropriate. While selectivity problems most routinely occur at the low end of an assay range, it may be prudent also to evaluate selectivity at higher analyte concentrations. In cases where interference is concentration dependent, it is essential to establish the concentrations below which interference is possible. It may be necessary to adjust the lower level of quantification accordingly, before assay validation.	
331	14	**Proposed change:** Propose to rephrase to: *"Specificity is validated by using at least 10 sources of sample matrix, spiked at ≤ 3-fold LLOQ and at about 75% ULOQ".*	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
331	23	IDEM 4.1.8 This is difficult to apply to pre-clinical animal matrices.	This is acknowledged, however the same criteria applies (10 sources).
331	36	Clarity is requested around wording on sources of sample matrix. Is this 10 individual naive matrix samples or ? What is the agency thinking on the use of 10 as opposed to any other number. Clarity also sought on the term, "near LLOQ". Is a spike at 3X LLOQ (typically low QC) acceptable? Or should spike be between LLOQ and LQC? **Proposed change :** Greater clarification of specific requirements is requested	The text has been revised. 10 independently lots from different individuals/animals should be used. Pooling of matrix is not acceptable.
331	45	We recommend spiking at 2 times the LLOQ and setting a criteria that 8 of 10 must recover within 75 -125% of nominal. Preparing 10 spikes at the LLOQ could result	There is no consensus about criteria to be applied.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		in several of the spike recovering just below the LLOQ and would not quantify (unless you are extrapolating).	
331,335	47	Comment: Regarding the number of samples used for validation for specificity in line 331. A population set of n=10 is applicable to ligand-binding assays (LBAs) and LC/MS methods. ADA methods do not appear to be addressed. Regarding the acceptance criteria in line 335. The acceptance criteria may not be applicable to LBAs. **Proposed:** Recommend that acceptance criteria that are applicable to LBAs be clarified, as well as the number of samples for ADA (the 2009 FDA draft guidance suggests 100, although many in industry are using 50).	ADA is not in the scope of this guideline. The text is revised and a complete new and separate section has been introduced on ligand binding assays.
331 - 340	24	Comment: QC acceptance ±20% (±25% at the LLOQ and ULOQ). This can be considered to be the currently accepted industry standard. This chapter should be aligned with our comments to lines 118 to 123. For some analytes there may be endogenous binding proteins, soluble receptors and other molecules that could bind to the analyte, this may be considered.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
332	3	Add "drug target and/or" to sentence "The presence of endogenous protein, drug target and/or endogenous antibodies to the analyte…"	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
334-335	21	To be consistent with line 349, replace may exceed 20% of the LLOQ by should not exceed 25% of the LLOQ. **Proposed change (if any)**: To replace "Due to interference of endogenous compounds, the blank response may exceed 20% of the LLOQ" by "Due to interference of endogenous compounds, the blank response should not exceed 25% of the LLOQ".	This is a different issue; it deals with selectivity testing.
334-335	23	To be homogeneous with line 349, replace "may exceed	This is a different issue; it deals with selectivity testing.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		20% of the LLOQ" by "should not exceed 25% of the LLOQ".	
334-335	36	**Comments;** Clarification on how much above 20% LLOQ is acceptable. Suggested to provide examples of how sponsor should document no impact on assay accuracy	There is no consensus about criteria to be applied.
339		**Proposed change (if any):** Suggest that more detail should be added on how this (parallelism experiment) should be performed.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
339 - 340	3	Add more detail on how this experiment should be performed.	This is not appropriate for a guideline, but for a SOP.
339 - 340	3	Analysis of 4PL equation fits are generally not included in the assay acceptance criteria. It would be more appropriate to discuss recovery of QC samples across sample and standard matrices.	This is agreed. The text is revised and a complete new and separate section has been introduced on ligand binding assays.
339 - 340	36	Analysis of 4PL equation fit are generally not included in the assay acceptance criteria. It would be more appropriate to discuss recovery of QC samples across sample and standard matrixes. **Proposed change** Recovery of QCs should be discussed rather than parameters generated for the standard curve fit (e.g. slope and asymptotes).	This is agreed. The text is revised and a complete new and separate section has been introduced on ligand binding assays.
339-340	45	The regression parameters generally are not highly reproducible. This sentence should say that QCs (spike) recovery should be unaffected.	This is agreed. The text is revised and a complete new and separate section has been introduced on ligand binding assays.
341 - 342	36	Statement regarding durable and stable antibody/antigen complex formation is very unclear. In particular clarity on what is meant by durable and stable binding? For example , Antibody / antigen complex may be stable in some and not at other conditions. **Proposed change** Clarification is requested or statement regarding durable	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		and stable antibody / antigen binding should be removed.	
341 - 344	3	It is not clear how a reference standard should be selected "in such a manner that specificity is ensured, and binding characteristics are durable…" There is no way to ensure selectivity. Reference standards are not selected to ensure stable binding; that is the challenge of the critical reagents (binding antibody pairs, capture/detection reagents etc). Certainly a change in reference standard as development proceeds is likely to provide purer material with different binding characteristics. We suggest that the wording be changed to indicate that the purest reference standard available at the time should be procured. There also needs to be some mention that the reference standard material used for the assay is comparable to the drug to be administered to animals or humans. The discussion for reference standard may include the fact that many protein biotherapeutics are produced in cells and thus, just like in the body, have a range of acceptable and active forms.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
345 - 349	3	Since LBA do not automatically revert to a single linear curve fit, some discussion on the selection of the "best fit" of a calibration curve (using weighting, etc.) is warranted. Unlike small molecules, LBAs are not developed to bracket the expected range of the samples. In fact, dosage of biotherapeutics frequently occurs in the mg/kg range and samples must be diluted. This fact overlaps into handling of QCs. Note the differences from small molecule in that for protein drugs: • The quantifiable curve range is not modified to meet the expected drug concentrations. Instead, the samples are diluted into the quantifiable range; • The concentration-response curve is not linear; rather it's sigmoidal with anchor calibrators included	Selection of the best fit is considered appropriate for a SOP, and not for a guideline. The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		< LLOQ and > ULOQ to improve cure-fit quality. • Response parameters are not used as acceptance criteria (at best they are used during development and sample analysis to monitor trending etc, but no criteria are typically applied). In addition, any changes to the regression model after method validation should require partial validation. Accuracy is discussed, but LBA typically use replicate samples (duplicates or triplicates). Therefore, the Coefficient of Variation (CV) should also be discussed, as closeness of fit of replicates adds value to the method. In the discussion of anchor points further discussion is needed on not applying acceptance criteria, and not reporting values for samples below the defined ULOQ. Some suggested wording: **Standard Curve performance** –at least 6 non-zero calibrators (typically in replicate) are needed to fit the curve. Standard calibrators should be spaced approximately evenly on a logarithmic scale within the anticipated range. Anchor points may be used to help define the curve, but do not require acceptance criteria since they are beyond the quantifiable range of the curve. All runs conducted during the validation must be reported in a table to establish the overall robustness of the method, with a minimum of 6 required. The regression model established during method development should be confirmed in a minimum of 6 independent pre-study validation runs, typically in the same runs in which method precision and accuracy are assessed. Model confirmation should precede the reporting of analytical results for validation samples. A standard point may be edited from a curve due to a technical error with an assignable cause (e.g., pipetting error) or through the application of *a priori* statistical criteria.	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		1.0 Accuracy: *A minimum of 75% of the individual non-zero standards must have a mean back calculated concentration equal to or within 20% of the nominal value, except at the LLOQ and ULOQ where the mean back calculated concentration can be equal to or within 25%.* 2.0 *Precision: While precision is not assessed during LC-MS, LBAs are typically run in replicates. Therefore, it may be useful to place acceptance criteria on the performance of the replicates. The %CV between each standard replicate must be equal to or within 20%, except at the LLOQ and ULOQ where the %CV can be equal to or within 25%. Specify criteria should be for measured responses or back-calculated values.* *3.0 At the end of validation, the cumulative mean %RE and percent CV from all runs should be calculated for each calibrator. The regression model is acceptable if both the RE and CV are equal to or within15% for each standard calibrator, not including anchor points, except at the LLOQ where both should be equal to or within 20%.*	
345 - 349	24	Comment: **Proposed change (if any)**: Please rephrase "In case anchor calibration standards are used, the additive value should be clearly documented": to: Once the model used for regression including or excluding anchor point is selected, it has to be used for the validations and routine analysis.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
345-349	47	Comment: Regarding the calibration curves. **Proposed:** Recommend clarification that anchor standards should not be included in the calculation of the calibration curve	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		and should not be used for acceptance or rejection of the curve. Referring to them as "anchor calibration standards" within the context of the guidance implies that they are part of the calibration curve. Furthermore, it should be stated that anchor points should not be removed to make QC samples pass, and if there are acceptance criteria for anchor points, like calibration curve points, the order of rejection should be clarified. Since anchor points are recommended to ensure consistent curve fitting for non-linear slopes, the question as to whether anchor points can be removed to allow for calibration curve points to pass could not be concluded, other than it should be documented a priori. Recommend that the EMEA considers this question.	
347	3	Anchor points. It is more appropriate to understand the non-linear function of standard calibrator, and the utility of anchor points in facilitating the curve fitting, than to require documentation of the additive value of anchor points. The use of anchor points is determined during method development and confirmed during method validation. Anchor points should not have acceptance criteria if anchor points are used.	
347	9	The use of anchor points are widely accepted for ligand binding assays; their use is usually determined during method development and confirmed during method validation. Anchor points should not have acceptance criteria if used, as these criteria are applied to calibration samples within the assay range. **Proposed change (if any):** Suggest to remove the need to justify their inclusion.	This is agreed.
347	36	**Comments**: Clarification on how sponsor should demonstrate value of anchor points is requested.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
347-348	21	Please provide clarity as to why it is necessary to clearly document anchor points, as the use of anchor points is inherent to 4-PL and 5-PL regression **Proposed change (if any)**: Delete "In case anchor calibration standards are used, the additive value should be clearly documented."	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
347-348	45	Please provide clarity as to why it is necessary to clearly document anchor points	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
349	45	25% at ULOQ only in case the calibration curve shows saturation behavior.	This criterion seems not common.
350	47	Comment: Regarding the use of "samples." **Proposed:** Recommend that "levels" be used instead of "samples."	The text will be revised.
350	51	Comments: Interpretation is 3 total, not three at each of three levels (9) **Proposed change (if any):** Clarify that at least three QC samples at each of the QC levels…	It is meant, 3 at each level. The text will be revised.
350 - 354	3	During validation, the QCs are used to define and validate the quantifiable range of the standard curve. Since QCs represent the sample to be assayed by the method, they should be spiked into the same matrix in which the samples are collected. To assess accuracy and precision, each run will contain QCs at a minimum of five levels. The levels of QCs to be prepared are: LLOQ, Low QC (3x LLOQ), mid QC (50% of ULOQ), High QC (75-90% of ULOQ), and ULOQ. Accuracy: (% RE) *The target mean calculated result for each QC must be equal to or within 20% of the nominal value, except at the LLOQ and ULOQ where the mean calculated result can be equal to or within 25% of the*	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		nominal value. *Precision: (%CV) The target %CV of the replicates for each QC must be equal to or within 20%, except at the LLOQ and ULOQ where the %CV can be equal to or within 25%* *Assessment of QC Accuracy and Precision:* As described in detail in the DeSilva White Paper, the process to assesses inter- and two types of intra-assay accuracy and precision can be reported in one table for each QC level. In this example, each of the five QCs is assayed in duplicate or triplicate over 6 runs to assess the "with-in run" intra-batch and pooled intra-batch, as well as the overall inter-batch accuracy and precision. This method results in 5 individual tables, one for each QC level and documents both intra and inter- batch accuracy and precision in each table. Add discussion on Dilutional Linearity and Prozone (Hook) Effect: **Dilutional linearity and hook effect** – Due to the nature or format of many immunoassays, the range of quantification (LLOQ to ULOQ) of the standard curve may be very narrow. For this reason it is necessary to show that the analyte of interest, when present in concentrations above the range of quantification (above ULOQ), can be diluted to bring the analyte concentrations into the validated range for analysis by the method. An additional reason for conducting dilutional experiments is for the identification of a possible prozone or "hook effect".	
350-354	45	Remove absolute error since this parameter has no useful meaning.	The guideline does not state the term absolute error.
350-354	48	Comment: 1- The meaning of the sentences is not clear .	1/2/3. The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
350-354 and 513-514		2- Moreover why these supplementary acceptance criteria are only added for the ligand binding assay. 3- The tables which present all the results obtained for the QC during the study have they to include also the QC belongs to the within – and between run accuracy and precision assessment? (question relating to line 513-514 too) **Proposed change (if any):** <u>For point 1:</u> For each QC level, the sum of the absolute mean accuracy and the absolute mean precision calculated from the totality of the QC obtained during the study (*included QC coming from failed series?*) should be < 30% (<40% for LLOQ level)	
351 - 352	3	Total Error calculations are most appropriate to determine the robustness of the assay during *method development* and not as acceptance criteria during validation.	According to the White paper, this is considered as validation.
352	11	Comment: "mean accuracy" should be "mean bias". It is also necessary to be more specific for the total error, i.e. the sum of all replicates for each level or the sum of all replicates of all levels.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
353-354	47	Comment: Regarding the discussion of results of runs to be used for validation. **Proposed:** Recommend that where it reads that the results of all runs should be included for validation, clarify if the expectation is to report the QC sample data and statistics for all accepted runs or not.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
355	9	Comment: Further clarification on what is meant by a "run" is required to apply this comment to ligand binding assays	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
355	45	For BAs, intra plate variability and drift within a plate should be evaluated during validation. If reasonable variability is shown, then it is unnecessary to have QCs placed throughout the plate. This appears to be more relevant than for chromatography methods.	The comment is noted. It is normal practise that every plate consist calibration samples, QCS and study samples.
355 - 356	3	Clarify definition of run size. For example, the number of assay plates run *per analyst* in a single instance. Note that each plate has a standard curve and its own QCs for monitoring assay performance.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
355-356	14	Comment: We recommend deletion of the sentence, since the resulting batch sizes from clinical studies from clinical development drugs are not foreseeable during method validation. Thus, the robustness of a method can only be assessed from real studies and not with an artificial run. **Proposed change:** Please delete the sentence.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
355 - 356	24	Comment: **Proposed change (if any):** We recommend to delete this paragraph, since the resulting batch sizes from clinical studies from clinical development drugs cannot be foreseen during method validation. Thus, the robustness of a method can only be assessed from real studies and not with an artificial run.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
355-356	45	What is the purpose of this statement? the size of an analytical run for LBA is 1 titer-plate (made clear in lines 362-364).	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
356	45	**Proposed change (if any):** spelling error: it should be "vice versa"	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
357 - 360	36	Antibody binding affinity is only one example of the biological integrity of biological drug candidate. This should be an example (e.g.), not as a mandatory request (i.e. in the guidance text).	The text is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Also it is not clear how a BA method can evaluate chemical and physical properties of biopharmaceutical drugs. This statement is most appropriate for a small molecule or should be clarified. **Proposed change :** In the guidance text "…but also biological integrity (i.e. maintenance of antibody binding affinity)…" the "i.e." should be replaced with "e.g.". Also clarification on the statement regarding use of BA method to evaluate chemical and physical properties of drugs should be clarified or removed.	
357 - 361	3	This statement is most appropriate for small molecules. Binding differences can be evaluated for large molecules, but it is not clear how an LBA method can evaluate chemical and physical properties of biopharmaceutical drugs.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
357-361	21	This recommendation seems to be too extensive. Typically sample preparation is required to make analyte detection possible. In most cases, it is not possible to analyze the analyte (and characterize to the recommended level described in the guideline) before sample preparation. Affinity determination in biological matrix is clearly outside of typical scope of bioanalytical method validation **Proposed change (if any):** Replace lines 357 - 361 by: Impact of sample preparation and use of additives on the stability and biological integrity of the analyte should be considered during method development. In addition, this also applies for changes after validation, e.g. a change in used diluents or reagents during analysis of study samples.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
357 - 361	24	Comment: **Proposed change (if any)**: please replace the	Comment noted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		sentence by: "Impact of sample preparation and use of additives on the stability and biological integrity of the analyte should be considered during method development. In addition, this also applies for changes after validation, e.g. a change in used diluents or reagents during analysis of study samples."	Changes after validation need a partial validation. The text is revised and a complete new and separate section has been introduced on ligand binding assays.
358-360	45	Evaluation should not only address chemical and physical properties, but also biological integrity (i.e. maintenance of antibody binding affinity)" – this is vague in terms of what antibody is referred to and what may be the practical implementation of this requirement since affinity determination in biological matrix is clearly outside of typical scope of bioanalytical method validation.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
362 - 364	3	Move to Section 5. Analysis of Study Samples.	The text is revised and a complete new and separate section has been introduced on ligand binding assays.
365	47	Comment: Regarding commercial kits. Many commercial kits come 'qualified' for an intended use, but many laboratories will nonetheless use the kits as a 'validated' method. Additionally, commercial kits originally developed and qualified for human matrices are often used in non-clinical studies for non-human matrices and should require further validation to demonstrate that they are suitable for use in these matrices. **Proposed:** Recommend that the validation of a kit be defined, particularly if the EMEA wishes to distinguish this differently than a method. Additionally, recommend including definitions for validation and qualification in the Definitions section, especially in light of validation/qualification activities for urine matrices, tissues, and biomarkers (now utilizing both immunoassays and LC-MS/MS techniques).	The text has been revised.
365 - 369	3	Commercial kits need to be fully validated to the same degree as internal assays. This includes a discussion on	This is agreed. The sentence has been changed with reference to the LBA section.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		preparation of an appropriate levels and matrix of QCs, calibration curve and appropriate curve fitting (i.e., not linear).	
365-369	21	Use LLOQ instead of LOQ for consistency	The text has been revised.
365 - 369	24	Comment: **Proposed change (if any):** Please use LLOQ instead of LOQ	The text has been revised.
365-369	26	Line 368: Is "LOQ" intended to mean "LLOQ"? Lines 367-369: Are accuracy and precision of the QCs the only requirement for validation of a commercial kit? What about purity and integrity of calibration standards (which can vary greatly), specificity, etc.? **Proposed change (if any):** Add additional text , as appropriate	Sentence has been changed with reference to the LBA section.
365-369	45	Use LLOQ instead of LOQ	The text is revised.
366-367	47	Comment: Regarding commercial kits. If a commercial kit is not initially validated by the vendor, it cannot be re-validated, as the wording suggests. Additionally, a kit is used for its materials/reagents during a method validation, and this should be clarified. **Proposed:** Recommend clarifying that all of the principles for a method validation apply when using a kit for reagents. Note that many kits do not contain matrix QC samples, but do contain controls. Matrix QC samples therefore need to be prepared, qualified, and validated, as with any method validation. Recommend that this be clarified in the guidance. Commercial kits usually come with expiration dates for	Sentence has been changed with reference to the LBA section.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		its reagents. The earliest expiration date of a kit's reagents should not apply to all reagents, especially reagents that impact a method, and would require a partial validation or at least a qualification of the reagent before continued use of the method, such as an antibody, conjugated materials, blocking buffers, and coated plates. This should also apply to changes in lot numbers. Recommend that this be clarified in the guidance.	
367	11	Comment: Is the example an intended use or other use?	The text has been revised and the example is deleted.
368	3	Change "LOQ" to "LLOQ."	The text is revised.
368,568	9	**Proposed change (if any):** Correct LOQ to LLOQ (Correction should be made for all occurrences in the document).	The text is revised.
370 - 374	24	Comment: **Proposed change (if any):** We would propose to EMA to replace "After complete validation of the analytical method, analysis of study or subject samples may be carried out ..." by "Analysis of study or subject samples can only be performed after validation of the method. Parameters that should be validated before starting the analysis of study samples should be defined in a SOP" We would propose to EMA to replace "Depending on the time period between validation and the analysis of the study samples it may be necessary to verify the performance of the method before start of the analysis of study samples" by "When relevant, the way method performance is verified before quantifying study samples should be described in an SOP".	Parameters that should be evaluated may indicate that not all parameters as described in this guidance should be taken into account. This would actually mean that no complete validation. has to be carried out. This is not agreed. The proposed change is considered to vague.
370-377	2	Comment: • The Guideline indicates that 'After complete	The text on long term stability has been revised. The guideline

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		validation of the analytical method, analysis of study or subject samples may be carried out'. Due to the broad definition of validation presented previously in the Guideline, this would suggest that long term stability in matrix must be complete. This is not practical as long term stability testing is an ongoing practice that may generate data over many years. • As for the previous section "Method validation" it should be clearly stated what are general applicable parameters for all bioanalytical methods and clearly described the specific requirements for the specific methods (LC-MS, ligand binding) in separate sections.	indicates now that results should be available before the study report is issued. The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
370 - 377	3	It would be helpful for clarity if sample analysis of small molecules followed the small molecule validation section and a specific sample analysis of biotherapeutics followed the LBA section. There are enough distinctions between the two processes to justify this organization. • LBAs typically do not extract, they use replicates for samples, as well as QCs. • LBAs are often plate based with specific recommendations for that scenario. • Dilution is usually required. • Critical reagents have lot to lot challenges.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
370-377	21	Wording: Replace "complete" by "full" validation. One cannot afford running a study after all parameters relating to a full validation are completed (including ISR, long term stability)	The text has been revised and states now 'full validation'.
370-377	45	Wording: Replace "complete" by "full" validation. One cannot afford running a study after all parameters	The text has been revised and states now 'full validation'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		relating to a full validation are completed (including ISR, long term stability)	
370-466	23	**Comments**: Methods for Analysis of study samples may vary due to the large amount of analyte types. **Proposed change (if any):** Review the wordings so the text can accommodate the different possibilities.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
371	16	Comment: The Guideline indicates that 'After complete validation of the analytical method, analysis of study or subject samples may be carried out.' Due to the broad definition of validation presented previously in the Guideline, this would suggest that long term stability in matrix must be complete. This is not practical as long term stability testing is an ongoing investigation that may generate data over many years. **Proposed change (if any):** Please permit sample analysis to commence upon completion of the necessary accuracy and precision runs and appropriate short term stability testing with the expectation that additional long term storage stability data in the matrix will be gathered during development and be sufficient to cover the duration and storage conditions experienced by study samples. Such additional stability data to be included as addenda to the validation method report.	The text on long term stability has been revised. The guideline indicates now that results should be available before the study report is issued.
371-372	23	To define what is a complete validation of the analytical method. **Proposed Change:** To replace "After complete validation of the analytical method" by "After the characterization of the main parameters such as within and between parameters, LLOQ, ULOQ, and matrix effect".	The text has been revised and states now 'full validation'. It is considered that the meaning of this is sufficient clear.
371-374	14	Comment: A complete validation would not include long-term stability data since these data might be still under investigation during study sample analysis. Moreover, it	The validation of the bioanalytical method should have been completed before analysis of study samples, except for long term stability. It should be ensured that the bioanalytical is accurate and reliable before the start of study sample analysis

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		is the risk of the sponsor not to investigate all validation parameters (e.g. matrix effect especially in case a stable-labelled internal standard is available) before analysis of study samples. Depending on decision processes within a company, parameters could also be investigated afterwards. Hopefully, not for every animal species a full validation is required (see also comment on line 86). Obviously when studies are run directly following one after another you could (per SOP) skip to verify the performance of the method before start of the analysis of study samples. **Proposed change:** Please delete the word "*complete*": "*After validation of the analytical method, analysis of study or subject samples may be carried out*."	and this can only be proven by completion of a full validation. The text on long term stability has been revised. The guideline indicates now that results should be available before the study report is issued. The text has been revised and the statement on 'time period between validation and study sample analysis' have been deleted for clarity.
372	8	Comment: The statement "Depending on the time period…" is very vague. After what period of time is a performance verification necessary?	The text has been revised and the statement on 'time period between validation and study sample analysis' have been deleted for clarity.
372	23	When do we need to verify the performance of the analytical method before we run sample analysis? Could it be possible to clarify "time period"?	The text has been revised and the statement on 'time period between validation and study sample analysis' have been deleted for clarity.
372-373	36	Clarification is sought around the agencies thinking on what is a suitable time period between method validation and analysis of study samples that precludes the necessity to verify assay performance.	The text has been revised and the statement on 'time period between validation and study sample analysis' have been deleted for clarity.
372-374	50	Depending on the time period between validation and the analysis of study samples it may be necessary to verify the performance of the method before start of the analysis of study samples The performance frequency and experimental design should be incorporated to this section to avoid any ambiguity.	The text has been revised and the statement on 'time period between validation and study sample analysis' have been deleted for clarity.
375 - 377	24	Comment:	Agreed on proposed change. The text has been revised

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): We would propose to EMA to replace "The study samples should be processed in accordance with the validated analytical method, and together with the QC samples and the calibration curve to ensure the acceptability of the analytical run" by "The study samples, QC samples and calibration standards should be processed in accordance with the validated analytical method to ensure the acceptability of the analytical run ".	accordingly.
378,317	9	The document does not mention duplicate analysis of study samples for ligand binding assays, or the need to have additional criteria for data acceptance based on comparison of these duplicates. This should be addressed as samples are normally run in duplicate in ligand binding assays.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
378-382	21	This section requires separating out LBA and LC-MS based assays f.i. a blank sample is considered as standard zero for ligand binding assays. **Proposed change (if any):** Replace "An analytical run consists of the blank sample (processed matrix sample without analyte and without IS) and a zero sample (processed matrix with IS)" by "An analytical run consists of the blank sample (processed matrix sample without analyte and without IS) and a zero sample (processed matrix with IS) for physico-chemical assays."	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
378-382	45	1. Blank plus IS sample (zero sample) not applicable for ELISA. 2. Add batch definition for ELISA (each plate one batch).	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
378-395	2	Comment: • The guideline should not specify where the Standards and QC samples should be placed. There should be enough scientific knowledge in	The text has been revised and states now: 'The QC samples should be divided over the run in such away that the accuracy and precision of the whole run is ensured.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		the laboratories to decide on this based on scientific reasoning. However, it's important that the study samples in a run are embraced by either Standards and/or ideally QCs. • If within run (Intra) precision and accuracy is to be reported on routine batches as indicated in the reporting requirements line 513 & 514 then at least three QC's per-level are required. • Line 381: If within run (Intra) precision and accuracy is to be reported on routine batches then at least three QC's per-level are required. • Line 392-395: QC samples should be distributed throughout the run. The restriction proposed is not justified scientifically.	This is only applicable to validation. The text has been revised and this is not specified anymore. The text has been revised and states now: 'The QC samples should be divided over the run in such away that the accuracy and precision of the whole run is ensured.
378 - 395	3	The document does not mention duplicate analysis of study samples (replicates) in large molecule assays; this should really be added. Also worth adding is the use of non-specific binding sample (where appropriate) and that large molecule bioanalysis typically does not perform sample processing.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
378-395	5	Comment: An analytical run consists of the standard curve including blank and zero sample, the QCs (at least 6 (2 low, 2 medium and 2 high) or at least 5 % of the number of study samples) and a defined number of study samples. QCs are used for the decision of acceptance/rejection of the analytical run. Neither the draft EMA guideline nor the FDA guidance clearly states whether or not the QCs used for the analytical run acceptance should be QCs, which are prepared freshly or should be "stored" frozen QCs, which are stored under the same conditions as the study samples. According to our experience both approaches	The guideline does not indicate that one of the methods should be used. Both methods are considered acceptable.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		are used in practice. Some companies prefer "fresh" QCs, other companies prefer "stored" QCs. It should be clarified whether one or both procedures are acceptable. It should be considered that results of "stored" QCs can be a mix of stability investigation and run acceptance/rejection. Therefore, for acceptance/rejection of analytical runs freshly prepared QCs should be preferred.	
378-395	26	Line 393: Seems contradictory to state that the QCs must be at beginning, middle, and end of run when the following sentence states these are only needed at beginning and end. QCs should be at beginning, middle, and end of run to perform their sentinel activity most effectively. Line 395: Would like to see dilution QCs in every run involving diluted samples to verify dilution performed correctly for that run.	The text has been revised and states now: 'The QC samples should be divided over the run in such away that the accuracy and precision of the whole run is ensured. Dilution integrity evaluations should be carried out during partial validation and not during study sample analysis.
378-395	45	This section provides restrictions on how to place QC samples in the run. If there is no carryover, the QCs could also be spread out randomly amongst the samples.	The text has been revised and states now: 'The QC samples should be divided over the run in such away that the accuracy and precision of the whole run is ensured.
378 - 398	24	Comment: Please adapt this paragraph to LBA. E.g., Blank plus IS sample (zero sample) not applicable for LBA, please define analytical run for LBAs. **Proposed change (if any):** Please replace "... should be processed *and extracted* as one single batch of samples" by "... should be processed as one single batch of samples" Please clarify / homogenize terms "full analytical run" and "individual batch of samples"	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
379-380	11	Comment: Only one blank and one zero sample in an analytical run?	More may be added; this is the minimum criteria.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
379-380	23	What criteria should be used for blanks (double blank) and zero sample (single blank) during validation and study assays? How shall the high, and mid QC concentrations be handled during analysis of study samples? For ligand binding, a blank sample is considered as standard zero. **Proposed Change:** To replace "An analytical run consists of the blank sample (processed matrix sample without analyte and without IS)" and a zero sample (processed matrix with IS)" by "For physico-chemical assays, an analytical run consists of the blank sample (processed matrix sample without analyte and without IS)" and a zero sample (processed matrix with IS)".	The matrix used for calibration samples and QC samples should be free of interference. During validation the cut off is less than 20% of the LLOQ. In case during subject sample analysis this value is exceeded, this should trigger questions/investigations. It is considered that this should be a SOP issue and not a guideline issue. The comment regarding QC samples is not understood. The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
379-395	19	Dilution QC samples should also be included in the analytical run when the samples are analysed with dilution in order to verify the dilution procedure. **Proposed change (if any):** We suggest to include 3 dilution QC samples per dilution procedure in the run.	Dilution integrity evaluations should be carried out during validation or partial validation and not during study sample analysis.
380 - 381	20	Comment: The phrase "…a set of calibration standards at a minimum of 6 concentration levels…" is somewhat redundant since this criteria was in place for MV and should therefore already be met. Since the calibration curve should not change from MV to sample analysis, a set of calibration standards as specified in the respective method should be included **Proposed change (if any):** Recommend revision to read that a set of calibration standards as specified in the respective method should be included in each batch	It is considered that the text of the guideline is sufficient clear.
382	23	What is "one single batch"? How will be defined the acceptance criteria for each batch and for the full run?	A batch is comprised of samples which are handled at the same time, i.e. subsequently processed without interruption in time and by the same analyst with the same reagents under homogenous conditions. Acceptance criteria should be applied

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			for the whole run and the separate batches in the run. The text has been revised for clarity. Line 386: preferably has been deleted. And in section 5.1 it will be added: If such an approach cannot be avoided (a single batch approach), for instance due to bench-top stability limitations, each batch of samples should include low, medium and high QC samples. Acceptance criteria should be pre-established in a Standard Operating Procedure (SOP) or in the study plan and should be defined for the whole analytical run and the separate batches in the run.'
382-389	48	Comment: Do you consider a manual extraction performed 46 samples by 46 samples (for example) as different batches even if each manual extraction is performed on the row (taking into account that the sample preparation performed before was conducted in the same time)?	A batch is comprised of samples which are handled at the same time, i.e. subsequently processed without interruption in time and by the same analyst with the same reagents under homogenous conditions. And in section 5.1 it will be added: If such an approach cannot be avoided (a single batch approach), for instance due to bench-top stability limitations, each batch of samples should include low, medium and high QC samples. Acceptance criteria should be pre-established in a Standard Operating Procedure (SOP) or in the study plan and should be defined for the whole analytical run and the separate batches in the run.'
384	23	How should we interpret the differences between "analytical runs" and "prepared runs". It is sometimes very difficult to prepare all samples for an analytical run at exactly the same time.	A batch is comprised of samples which are handled at the same time, i.e. subsequently processed without interruption in time and by the same analyst with the same reagents under homogenous conditions. And in section 5.1 it will be added: If such an approach cannot be avoided (a single batch approach), for instance due to bench-top stability limitations,

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			each batch of samples should include low, medium and high QC samples. Acceptance criteria should be pre-established in a Standard Operating Procedure (SOP) or in the study plan and should be defined for the whole analytical run and the separate batches in the run.'
384-385	45	Remove sentence:" Analysing as a single analytical run samples prepared separately as several batches should be avoided. " redundant to the sentence before	The text is included to ensure that it is avoided.
388 398 491	47	Comment: Regarding the reference to a "study plan" and/or "protocol." **Proposed:** Recommend consistent use of terminology within the guidance, particularly in light of definitions specific to FDA GLP regulations vs. OECD GLP principles.	In line 397, the 'analytical study protocol' has been changed into 'protocol. 'Line 491 has been revised.
390-391	21	Requirement is scientifically not sound. What about cross-over studies or long-term toxicity studies, where samples from the same subject are measured at different times? The sentence may be applicable only to special clinical studies (such as bioequivalence) If the method is valid and ISR has been demonstrated, then it shouldn't make a difference if samples of subjects are analyzed in different runs **Proposed change (if any):** Start line 390 with "Although other approaches may be applied, it is advised….."	The text has been revised and indicates now that this is specific for bioequivalence studies.
390-391	32	It is advised where possible or appropriate to analyse all samples of one subject etc – ie add "where possible" or "appropriate".	The text has been revised and indicates now that this is specific for bioequivalence studies.
390-391	34	Comment: Assays to test immunogenicity of drug product are performed at different times (for example 3 months and 6 months after the first drug injection)	Immunogenicity is not in the scope of this guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change: Please add in the text	
390-391	45	1. Requirement scientifically not sound. What about crossover studies or long-term toxicity studies, where samples from the same subject are measured at different times? The sentence may be applicable only to special clinical studies (such as bioequivalence) 2. BE study specific, if the method is valid and ISR has been demonstrated, then it shouldn't make a difference if samples of subjects are analyzed in different runs	1. The text has been revised and indicates now that this is specific for bioequivalence studies. 2. This is agreed. The text has been revised.
390 - 392	3	Analysis of all samples for one patient in one analytical run very often is not possible, or will require long storage of initially collected samples, e.g. for long clinical and non-clinical (e.g. long term toxicology) studies. In the case of LBAs, variability of the assay is controlled by evaluating the standard and QC performance on each assay plate and should allow for normalization and monitoring of assay performance and reported sample concentrations. The statement regarding analysis of all samples from a given patient should be either removed or changed to become a recommendation for some cases, e.g. when studies are short or assay variability is high(er).	The text has been revised and indicates now that this is specific for bioequivalence studies.
390-392	14	Comment: This should only be an advice for BE studies. For other studies with multiple periods (e.g. ascending dose) or lengthy late phase studies it is not feasible to analyse all samples of one subject in one analytical run. According to safety requests on dosing of subjects before dosing subjects with a higher dose, it is often required to have bioanalytical data earlier available as after the last sample has been taken. Furthermore, variability in pharmacokinetic data in those studies with a parallel group design is rather based in biological variability than in variability of bioanalytics. It is common sense that study samples of more than	The text has been revised and indicates now that this is specific for bioequivalence studies. In accordance with the proposed change, the second sentence has been deleted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		one subject are analysed within one analytical run, otherwise assay throughput will be directed to "ad absurdum". **Proposed change:** "*For bioequivalence studies it is advised to analyse all samples of one subject together in one analytical run to reduce the intra-subject variability in outcome."* Please delete the sentence: "*Moreover, it is considered acceptable to analyse study samples of more than one subject in one analytical run*."	
390-392	15	This should only be applicable for BE studies.	The text has been revised and indicates now that this is specific for bioequivalence studies.
390 - 392	36	**Comments**: Analysis of all samples for one patient in one analytical run may only be possible in certain studies (such as human BE studies). The driver for analysis is the known storage stability duration and for long clinical and non-clinical (e.g. long term tox) studies this may be an overriding factor. **Proposed change** Statement regarding analysis of all samples from a given patient should be either removed or better defined to indicate that this is prefer when study design and stability data allows.	The text has been revised and indicates now that this is specific for bioequivalence studies.
392-393	21	Position of QC samples is not applicable for ELISA.	This is agreed. The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
392-393	40	Comment: The guideline suggests placing the QCs at the beginning, middle and end of the analytical run. It would be better to have the QCs randomly distributed. **Proposed change (if any):** It is recommended that QCs should be randomly distributed across the run.	The text has been revised and states now: 'The QC samples should be divided over the run in such away that the accuracy and precision of the whole run is ensured.
392-393	45	Position of QC samples is not applicable for ELISA	This is agreed. The text on LBA is revised and a complete new

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
			and separate section has been introduced on ligand binding assays.
392 - 395	3	The statement regarding positioning of QC samples across a run is very LC-MS centric. In LBA methods, each plate has a separate set of standards and QCs.	This is agreed. The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
392-395	14	Comment: It should be possible to have different options on how to distribute the QCs over the run. The proposed position of QC samples is not applicable for ligand binding assays.	The text has been revised and states now: 'The QC samples should be divided over the run in such away that the accuracy and precision of the whole run is ensured.
393-395	23	The re-grouping of QCs at the beginning and end of runs appears acceptable in this document although this is not recommended by the FDA. Depending on the size of the run, it should specified that it is necessary to monitor the middle section.	The text has been revised and states now: 'The QC samples should be divided over the run in such away that the accuracy and precision of the whole run is ensured.
393-395	32	This sentence conflicts with the previous sentence and as such is largely redundant and should be removed.	The text has been revised and states now: 'The QC samples should be divided over the run in such away that the accuracy and precision of the whole run is ensured.
393-395	49	This is not acceptable especially for LC-MS/MS analysis. During LC-MS/MS analysis there is a possibility of instrument drift. This could happen at any stage during a run. If QC samples are at the beginning and the end of a run, there is no way to determine if instrument drift occurred during the batch. **Proposed change (if any):** QC samples should always be equally distributed throughout the batch.	The text has been revised and states now: 'The QC samples should be divided over the run in such away that the accuracy and precision of the whole run is ensured.
396 - 398	24	Comment: As a general comment, acceptance criteria for chromatographic and LBAs methods should be clearly defined in separate paragraphs	This is agreed. The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
396-416	43	Two very serious issues should be expressed to this paragraph:	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		I.: The criteria for the acceptance of the individual standards and QCs samples are based on their "accuracy" i.e. on the deviation of the individual from the nominal values. These deviations should be within ±15% except for LLOQ (lower limit of quantification), when it should be within ±20%.. However the same criteria are set up for the statistical terms precision and accuracy (mean accuracy) in the pre-study validation. Precision is defined here as the degree of scatter or probability of the deviation of the individual point from the mean value as determined by the relative standard deviation. This must be within ±15% (±20% for LLOQ), while accuracy (mean accuracy) defined here as a percentage of the mean value of the nominal within 85 - 115% (80 - 120% for LLOQ). This set up of the equal criteria for accuracy and precision as statistical terms and accuracy of the individual points is inconsistent. The inconsistence may be illustrated using the next two models: Model A: Accuracy and Precision of the method at all concentration levels are 100% and 15% respectively. According to the error distribution function approximately 68 % of all values are within the interval mean ± SD. ▪ If 6 QCs on three different concentration levels will be distributed within a run, the calculated probabilities that 2 QCs of the same concentration level or more than 2 QCs will be outside of 15% of the nominal (in this case mean) value may be presented as follows: 2 QCs of the same concentration level outside of 15% 6.57%	I. There is a general consensus within the scientific community and regulators for the use of the criteria as expressed in this guidance. It is recognized that other statistical methods may be applied, but all have their pros and cons.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome

3 QCs outside of 15%	20.61%
4 QCs outside of 15%	7.27%
5 QCs outside of 15%	1.37%
6 QCs outside of 15%	0.10%
Probability that the run will be rejected because of QCs	35.92%

This is a very high probability. If there are 24 runs, 8 of them should be rejected and analyzed repetitively. 2 of them will be rejected again and analyzed once more. There is a high probability that almost 50% of all batches have to be reanalyzed.

Model B: Accuracy and Precision of the method at all concentration levels are 110% and 7.5% respectively.

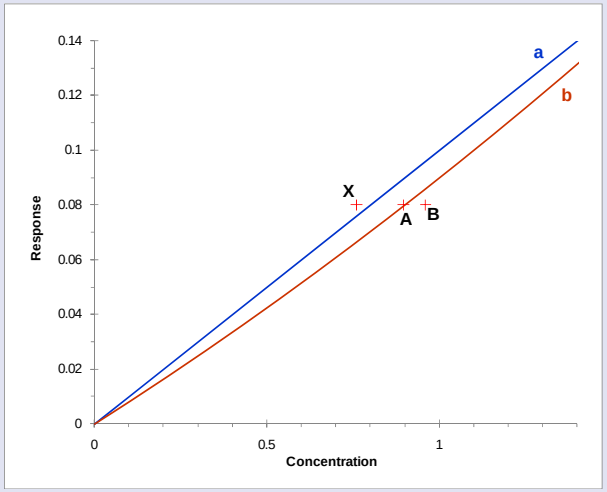
According to the error distribution function approximately 50 % of all values are within the interval mean ± 0.66*SD,

- If 6 QCs on three different concentration levels will be distributed within a run, the calculated probabilities that 2 QCs of the same concentration level or more than 2 QCs will be outside of 15% of the nominal value are presented:

2 QCs of the same concentration level outside of 15%	5.93%
3 QCs outside of 15%	13.18%
4 QCs outside of 15%	3.30%
5 QCs outside of 15%	0.44%
6 QCs outside of 15%	0.02%
Probability that the run will be rejected because of QCs	22.87%

This is still a very high probability as well.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		The inconsistency of the criteria set-up is obvious. Other models might be presented. This theoretical evaluation can be proven by many observations in practice whenever the real precision and accuracy of the method are approaching the values given in the models above. Suggestion: To solve this problem of inconsistency several solutions may be suggested e.g.: 1) By setting up the criteria loosely for the deviation of individual samples of standard curve and individual QCs during the process of routine drug analysis while keeping unchanged the criteria for precision and accuracy of the method in pre-study validation testing (20% for LLOQ and 15% for other concentration values of standards as well as QCs). For example 25% deviation for LLOQ from the nominal value and 20% deviation for other concentration values of standards as well as QCs might be adequate. 2) By keeping the criteria for the individual deviations from the nominal concentration values unchanged (20% for the LLOQ and 15% for other than LLOQ) and by tightening the criteria for precision and accuracy in the pre-study validation. E.g. the limits for precision and accuracy 12% for LLOQ and 8% for other concentration levels of standards as well as for all QCs might be adequate. II: It is to accept from the practical point of view that the exclusion of the individual points either calibration standards or QCs on the bases of their	II. Using the mean value is considered critical. First of all, the mean differs per run. Furthermore, a mean with a rather great bias may results in accepting QCs which are having a large deviation from there nominal value.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		difference from the nominal value is useful. However from the scientific point of view such an approach is basically not correct. As accuracy and precision are statistical terms based on mean values and standard error distribution around the means, the exclusion of the individual points should be done on the basis of their deviation of the mean value. There is a number of statistical tests for the exclusion of outliers from the set of the statistical data (e.g. Dixon test). All of them are based on the deviations from the means. Strictly speaking deviations limits 15% (20% for LLOQ) should be considered from means but not from nominal values. This may be illustrated by the figure below.  Let us consider the theoretical linear regression calibration curve (line) a. The accuracy (mean) in the whole calibration range is 100%. The real calibration curve b however is not an ideal line and	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		may be curved so that the mean accuracy in the range of medium concentrations (mean back calculated concentration values) is still within the limit 15%. Let s say the deviation of the point A from the nominal value here is 12%. Let us consider point B with the deviation from the nominal value + 20%. This point should be excluded according to the present guideline. However its deviation from the mean is 8% only and therefore it cannot be considered as an outlier. On the other hand point X with the deviation from the nominal value – 5% should not be excluded according to the present guideline while statistically it should be considered as an outlier (its deviation from the mean is 17%) Evidently the current approach may in some cases essentially "improve" (bias) the shape of the calibration curve. It is therefore to recommend checking potential points like B and X additionally after all data for evaluation of the means is available, whether inclusion or exclusion such points may be statistically justified.	
396 - 421	20	Comment: The role of blanks, zero standards and dilution QCs in batch acceptance are not included. Both blanks and zero standards must meet acceptance criteria for a batch to be accepted. Batch acceptance should not be based on the acceptability of the dilution QC samples, but the dilution QCs in the batch must be acceptable to report any diluted sample data in the batch **Proposed change (if any):** Recommend including acceptance criteria for blanks, zero standards and dilution QCs	The matrix used for calibration samples and QC samples should be free of interference. During validation the cut off is less than 20% of the LLOQ. In case during subject sample analysis this value is exceeded, this should trigger questions/investigations. It is considered that this should be a SOP issue and not a guideline issue. Dilution integrity evaluations should be carried out during validation or partial validation and not during study sample analysis.
396-429	2	Comment: • Not clear that this is only in reference to small molecule assays until lines 422-429 where LBAs are mentioned.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		• Acceptable standards may not be rejected to increase acceptance of QC results • Between run precision and accuracy are assay performance parameters which deserve reporting rather than serving as run acceptance criteria during study sample analysis **Proposed change (if any):** • Line 421: "precision should not exceed 15%". It should say either imprecision should not exceed 15% or precision should be within 15%.	This is agreed. - This is agreed. The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations. - Regarding the proposed change, the text has been revised for clarity.
396 - 429	3	Again elements that are LC-MS-specific need to be relocated to the LC-MS section, especially lines **407-416**. As an example of the confusion generated by combining the sections: since LBA nominally have a specific section in the present draft, what in lines **397-421** applies to LBA and what applies to LC-MS or both?	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
396 - 429	3	Section 5.2 Acceptance criteria of an analytical run is relevant for LC-MS methods. Include LBA specific criteria in LBA section.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
396-429	15	Each run should be valid by its own. **Proposed change:** The between-run accuracy and precision should be calculated as descriptive parameters.	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
396-429	21	Add ELISA acceptance criteria.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
396-429	45	Add ELISA acceptance criteria	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
396-429	47	Comment: Regarding acceptance criteria of an analytical run. Immunoassays and/or ligand-binding assays (LBAs)	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		acceptance criteria are not described here. Furthermore, there are additional LBA experiments that need to be defined and included for LBA. **Proposed:** Recommend that LBA acceptance criteria for applicable experiments need to be defined. The use of, and treatment therefore, should be similar to that of LC/MS assays, which is why LBA methodology is part of this draft guidance (as it should be). Refer to the rationale and recommendations described for other previous comments. Comment: These are defined later in this section for LBAs in terms of overall run acceptance in lines 422-429 although perhaps this section needs to be expanded to include the between run accuracy for LBAs, as well as all other applicable experiments and acceptance criteria, as discussed above.	
400-401	21	LLOQ values for ligand binding assay might be included. **Proposed change (if any):** To replace "The back calculated concentrations of the calibration standards should be within ±15% of the nominal value, except for the LLOQ for which it should be within ±20%" by "The back calculated concentrations of the calibration standards should be within ±15% (or 20% for ligand binding assay) of the nominal value, except for the LLOQ for which it should be within ±20% (or 25% for ligand binding assay)".	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
400-401	23	LLOQ values for ligand binding assays might be included. **Proposed Changes:** To replace "The back calculated concentrations of the calibration standards should be within 15% of the nominal value, except for the LLOQ for which it should be within 20%". By ""The back calculated concentrations of the calibration standards	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		should be within 15% of the nominal value (or 20% for ligand binding assays) , except for the LLOQ for which it should be within 20% (or 25% for ligand binding assays)".	
401-402	23	At least 75% of the calibration curves standards with a minimum of six, must fulfil this criterion». Does that mean that the curve is composed of 6 points and 4 are acceptable or do we need 6 acceptable points so at least 8 in the curve? Should it be clarified here that this relates to a minimum of 6 calibration samples at six different concentrations.....?	A comma is added to make it more clearer. At least 75% of the calibration curves standards, with a minimum of six, must fulfil this criterion. (this is also be applied to line 176) It means that a minimum of six samples should pass the criteria.
401 -402	36	**Comment** Clarity is suggested on the text to indicate that calibration standards out of acceptance criteria are to be deleted but the overall calibration must comprise of at least 75% of the total of the calibration standards (min of 6) that are within acceptance criteria.	A comma is added to make it more clearer. At least 75% of the calibration curves standards, with a minimum of six, must fulfil this criterion. (this is also be applied to line 176) It means that a minimum of six samples should pass the criteria.
403	49	Should be rejected or MUST be rejected. Can a standard be included in the curve if it is outside limits.	Must and should result in the same: rejection. Should is used in the guideline.
404-405	21	Criteria for decision should be pre-defined in a SOP or in the validation method or in study plan. **Proposed change (if any):** To replace "Criteria for decision to exclude calibration standards or not should be pre-defined in a SOP" by "Criteria for decision to exclude calibration standards or not should be pre-defined in a SOP or in the validation method or in study plan".	The text has been deleted, as it is already included in section 5.1 Analytical run.
404-405	23	**Proposed Changes:** To replace "Criteria for decision to exclude calibration standards or not, should be pre-defined in an SOP" by "Criteria for decision to exclude calibration standards or not should be pre-defined in an	The text has been deleted, as it is already included in section 5.1 Analytical run.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		SOP or in the validation plan or in the study plan.”	
407-410	8	Comment: The paragraph could imply that the acceptance criteria of the new LLOQ becomes ±20%. This should be clarified.	The lowest calibration standard in this case should have an acceptance of 15%. Otherwise a different accuracy/precision would be applied to study samples between the LLOQ and the next acceptable higher calibration standard in the other runs which do not have a failed LLOQ.
407-410	50	Comment: If the rejected calibration standard is the LLOQ, it should be realized that the LLOQ for this analytical run is the next acceptable higher calibration standard concentration of the calibration curve. If the highest calibration standard is rejected, the ULOQ, for this analytical run is the concentration of the next acceptable lower calibration standard of the calibration curve. In case of failure of LLOQ and ULOQ calibration standards, subsequent calibration standards can be accepted, if no significant deviation in slope of calibration curve observed or the integrity and performance of truncated calibration curve is established/proven during validation, should be incorporated to this section. As per our scientific understanding and practical experience, One LLOQ and/or One ULOQ can be excluded, if two Calibration curves are used otherwise we feel, no LLOQ CC standard or ULOQ CC standard can be excluded and the batch cannot be accepted.	As long as the same regression model is applied, it is considered that this will not have a large impact in general.
407-411	23	This is not in agreement with the FDA recommendation and current practices. It is not allowed to change the range of quantification. If the LLOQ or ULOQ is not acceptable then the run should be rejected. **Proposed Changes:** To replace “If the rejected calibration standard is LLOQ, it should be realised that the LLOQ for this analytical run is the next acceptable higher calibration standard concentration of the calibration curve. If the highest calibration standard, the	Comment noted. The EMA considers rejection not needed. As long as the same regression model is applied, it is considered that this will not have a large impact in general.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		ULOQ for this analytical run is the concentration of the next acceptable lower calibration standard concentration of the calibration curve." by " If the rejected calibration standard is the LLOQ or the ULOQ, the run should be rejected".	
407-411	37	Comment: This proposal is a non-scientific based. It is scientifically justified that two points of calibration curve are determined with higher uncertainty: LLOQ and ULOQ (see comment to lines 174-175). The chance to reject those points is therefore greater than for other points. The calibration range is set during method validation, and rejection of single calibration standard (including LLOQ and ULOQ) should not influence calibration range. Moreover, the proposal of shortening of calibration range is not consistent with described rejection of in-the-middle-calibration standards. In case of using the same logic for rejection of other calibration standards, there is a need to divide calibration range for 2 sub-ranges with immeasurable concentrations in the middle [e.g.: calibration curve points: 1,2,3,4,5,6; point 3 rejected, two sub ranges: 1-2 and 4-6; the concentrations in the range 2-4 are immeasurable]. This is strange proposal, which generates unnecessary work for both bioanalytical laboratory and the regulatory authorities. **Proposed change (if any):** If the rejected calibration standard is the LLOQ or ULOQ, the calibration range for analytical run is unaffected (the calibration range set during validation is used). ~~the next acceptable higher calibration standard concentration of the calibration curve. If the highest calibration standard is rejected, the ULOQ, for this analytical run is the concentration of the next acceptable lower calibration standard of the calibration curve. The revised calibration range must cover all QC samples (low, medium and high).~~	Extrapolation is considered not acceptable.
408	9	Suggest to clarify that, if duplicate curves are used and only one of the LLOQ or ULOQ standards fails, the range	The text has been revised according to the suggestion.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		is unchanged.	
412	9	This section needs to further clarify the acceptance criteria to be used if more than 6 QCs are run. If dilution QCs are also run, it would be useful to state the criteria for these.	The text has been revised and '4/6' has been deleted to prevent confusion.
412	21	Values for ligand binding assay might be included. **Proposed change (if any):** To replace "The accuracy values of the QC samples should be within ±15% of the nominal values." by "The accuracy values of the QC samples should be within ±15% of the nominal values or 20% for ligand binding assay."	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
412	23	Values for ligand binding assays should be included. **Proposed Change:** To replace "The accuracy values of the QC samples should be within 15% of the nominal values" by "The accuracy values of the QC samples should be within 15% of the nominal values, or 20% for ligand binding assays."	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
412-415	19	Acceptance criteria for dilution QC samples should also be included. **Proposed change (if any):** We suggest that the results of diluted study samples should be accepted if in a valid run for at least 2 of the 3 dilution QCs the measured concentration falls within nominal ±15%.	Dilution integrity evaluations should be carried out during validation or partial validation and not during study sample analysis.
412 - 415	24	Comment: **Proposed change (if any):** Please replace: "At least 4 out of 6 (67%) QC samples and at least 50% at each concentration level ..." by "At least 67% of all QCs and 50% at each concentration level ..."	The text has been revised and '4/6' has been deleted to prevent confusion.
412-416	1	The introduction in the study procedures of the within-run and between-run accuracy and precision off all sequences including all the obtained values is interesting	This is agreed, and considered logical. No addition in text considered needed.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		in order to put in evidence a global trend in the measurement (over-estimation, variability at a certain concentration range etc.). However abnormal values coming from several factors (e.g. incorrect internal standard addition) may lead to inappropriate estimates exceeding the proposed acceptance limits (± 15 %). **Proposed change:** It would more appropriate to exclude from calculation the results from abnormal QCs coming from documented and justified errors during preparation and/or analysis.	
413	23	**Proposed Change:** Change 50% to "at least one of".	This is not agreed, as this would only apply in case 2 QC samples at the same concentration level are used.
414-415	49	"....be reanalysed" Does this mean re-extracted or reinjected? If a reason for failure is identified (eg: instrument error) is it acceptable to reinject the batch?	The guideline clarifies this in line 460-463.
416	11	Comment: change "nominal value" to "respective nominal values"	The text has been revised and this recommendation is changed.
416	23	The between run accuracy should already have been validated and therefore is not useful here. Is the within-run precision calculated with all QCs (including the failed ones)? The QC samples rejected should be excluded from the calculation of the between run (mean) accuracy. The acceptance criteria on overall statistics of study QCs are not acceptable as study results are based on individual run acceptance.	This is agreed. The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
416	24	Comment: If each individual is successful according to acceptance criteria, but the overall precision and accuracy fall outside of overall acceptance criteria, how would EMA advice to proceed on the outcome of the study?	This would imply that the whole study should be rejected. However, the text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
416	45	Not a valuable acceptance criteria, carries the inherent	The text has been revised. Between run precision and

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		risk to make the entire study invalid by one run.	accuracy should be reported, however deviations may need additional investigations.
416, 421	8	Comment: The criteria outlined here are not present in the FDA guidance. What happens if, for a study, the criteria are not met (e.g. due to one outlier QC)? What is the expectation?	This would imply that the whole study should be rejected. However, the text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
416, 421	3	Are rejected QCs included in the statistics? (If yes, the impact could be quite significant). What is the consequence if the criterion is not met?	Rejected QCs based upon documented analytical reasons should not be included. If the criterion is not met, this would imply that the whole study should be rejected. However, the text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
416 and 421	4	Comment: The between run accuracy and precision should not exceed 15%. Each individual run has acceptance criteria for QC samples (lines 400 to 415). Results are reported, often during the course of a study, based on these acceptance criteria. It is unlikely, but possible, that every analysis run in a study could pass acceptance criteria for the QC samples within that run, but the overall (between-run) accuracy or precision could exceed 15% when all the QCs, whether they passed the acceptance criteria or not, are included in the summary statistical analysis. Obviously such an event would merit investigation and the precision should be improved, but how would the existing data in this situation be handled? Should data from individual runs be rejected so that the overall accuracy and precision meets the acceptance criteria, even though those runs passed the individual run acceptance criteria?	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Should all of the study samples be reanalysed with an improved method, even though many of these results could have been obtained from perfectly acceptable batches? Both of these options are undesirable. **Proposed change (if any):** Remove the acceptance criteria for between run accuracy and precision, or state the between-run accuracy of the QCs that met the acceptance criteria should not exceed 15%.	
416,421	9	It is not logical to apply additional between-run accuracy and precision criteria to run acceptance, if the specifications in lines 412-415 are already being applied to accept or reject individual batches. **Proposed change (if any):** Suggest to delete these sentences.	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
416 and 421	14	Comment: These are not valuable acceptance criteria, since each run is accepted on its own (within run) merit – see acceptance limits in above paragraph. These criteria carry the inherent risk to make the entire study invalid by one run. One may use these parameters only as descriptive parameters **Proposed change:** **"***As descriptive parameter the between-run accuracy should be calculated."*	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
416 and 421	21	Between run precision and accuracy are assay performance parameters which deserve reporting rather than serving as run acceptance criteria during study sample analysis **Proposed change (if any):** Delete lines 416 and 421.	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
416-421	23	Acceptance criteria on overall statistics of study QCs is not acceptable as study results are based on individual runs; acceptance, or criteria for overall statistics should be enlarged to be statistically compatible with the 15/4/6 rule.	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
416-421	29	Comment: The between-run (mean) accuracy of the QC samples should be within 15% of nominal value and between-run precision should not exceed 15%. This refers to global precision and accuracy of QC samples. While there is a criteria specified for within-run acceptance, ie at least 4 out of 6 (67%) of samples and at least 50% at each concentration level, must be within ±15% of nominal value, the between run acceptance is redundant. This is because, the value of between-run precision and accuracy is known at the end of the study when all the batches are completed and there is a possibility due to assignable or non-assignable causes, the precision, rather than accuracy, may fall out of 15%, where the whole study cannot be rejected. While these values of global precision and accuracy can be reported in the bioanalytical report, there should not be an acceptance criteria for between run precision and accuracy. When run acceptance (within run) is not met, the batch is rejected and study samples are reanalysed as specified in line numbers 414-415 of page 11/17. **Proposed change (if any):** The between run accuracy and precision of QC samples (with and without failed QCs) may be specified in the bioanalytical report, but does not form the basis of rejection of the batches.	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
416-421	32	Each batch stands on its own merits – this sets a second set of criteria which may conflict with the per batch acceptance and result in the entire study being invalidated on the basis of one run.	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
416-421	36	**Comment** Clarification is sought on this section. The agency	The text has been revised. Between run precision and accuracy should be reported, however deviations may need

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		appears to be introducing a requirement around between-run acceptance criteria. Current best practice does not use this as a run acceptance as each analytical run has individual run accented for the concentration data reported for that run. Wider accuracy and precisions criteria are also used in LBA analysis which are not detailed in this section.	additional investigations.
416 421	48	Comment: This section is relating to the acceptance criteria for an analytical run, however these lines are referred to the acceptance criteria of all the analytical runs. Except if these sentences mean that for each runs the mean between run accuracy and precision should be calculated? Moreover the tables presented the totality QC obtained during the study to calculate the between-run mean accuracy and precision should they include both the QC meeting and not meeting the acceptance criteria? Does they all include in the calculation even they belongs to not validated series due to too many QC outside of the acceptance criteria? **Proposed change (if any):** Add a section for the acceptance criteria of the assay study.	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations. QCs from rejected runs should not be included.
417 - 419	24	**Proposed change (if any):** data for the accepted analyte "should be" instead of "can be"	The guideline indicates 'can be' as it is up to the analist whether both analytes will be reanalysed or not.
417-419	50	Drug-drug reactivity and stability for multiple analytes for simultaneous estimation: It is suggested to prove no reactivity of each analyte(s) with other analytes and stability of each analyte(s) in presence of other analytes. This may ensure an extended and long term performance of a bioanalytical method.	The text on stability has been revised as follows: 'In case of a multi-analyte study and specific for bioequivalence studies, attention should be paid to stability of the analytes in the matrix containing all the analytes.'

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
418	8	Comment: In the statement "…the data for the accepted analyte can be used…", the use of the word "can" implies that we can choose to reject passing data, which is not good science. **Proposed change (if any):** Change the word "can" to "must".	The guideline indicates 'can be' as it is up to the analist whether both analytes will be reanalysed or not.
420 - 421	24	Comment: If each individual is successful according to acceptance criteria, but the overall precision and accuracy fall outside of overall acceptance criteria, how would the Agency advice to proceed on the outcome of the study?	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
420-421	45	1. precision should not exceed 15%". It should say either imprecision should not exceed 15% or precision should be within 15%. 2. Not a valuable acceptance criteria, carries the inherent risk to make the entire study invalid by one run	1/2. The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
421	23	The between run precision should already have been validated and therefore is not useful here. However the text states "The between run (mean) precision should not exceed 15%" The QC samples rejected (or at least the outliers) should be excluded from the calculation of the between run (mean) precision **Proposed Change:** Complete with "the between-run precision of the QCs samples should not exceed 15%" as in line 416.	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
421	23	The between run precision should already have been validated and therefore is not useful here.	The text has been revised. Between run precision and accuracy should be reported, however deviations may need additional investigations.
422 - 426	24	Comment: We would like ask EMA to use either an integrated approach (chromatographic and LBA methods) or separate chapters for each technique in the same discipline.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
423 - 426	3	Introduce duplicates, singling standards, etc.	The text has been revised for clarity.
423-426	45	1. Add ELISA acceptance criteria, add comment that ULOQ 25% is only valid in case the calibration curve shows saturation at the high end. 2. Suggest that standard curve acceptance criteria not be mandated but suggested. If appropriate methods are used to control the assay, e.g quality controls placed appropriately and masking of curve points are not allowed, it is suggested that this is sufficient control for the assay. This strategy has worked well in many LBA methods, specifically not allowing masking of curve points as this can be subjective unless there is a known technical reason for masking the calibrator/standard curve points	For LBA it is not requested to remove calibration standards. The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
423-429	48	Comment: The most ligand binding assays are performed in duplicate- However this draft document does not consider the CV% between a duplicate. What are the position of the authorities relating to this acceptance criteria → a CV% should be < 20% or 25% If the dulplicate have a CV > to the acceptance criteria, the whole duplicate should it be deleted from the calibration curve (for example) or the accurate simplicate could be considered and so maintained in the calibration curve? **Proposed change (if any):** Add acceptance criteria for the CV% between duplicate for standard calibration; QC and samples	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
427 - 429	24	Comment: **Proposed change (if any):** Please replace: "At least 4 out of 6 (67%) QC samples and at least 50% at each concentration level ..." by	The text has been revised and '4/6' has been deleted to prevent confusion.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		"At least 67% of all QCs and 50% at each concentration level ..."	
428	23	**Proposed Change:** Change 50% to "at least one of".	This is not agreed, as this would only apply in case 2 QC samples at the same concentration level are used.
430	23	Section 5.3 describes the possibility to extend the calibration range once having verified the precision and accuracy. It can also be necessary to verify the model of regression chosen.	The text has been revised. Response function has been added.
430-437	2	Comment: • Line 431: This is not typically done for LBAs as it is given that the assay range is small and that sample dilution will be required. • For the majority of LBA calibration range is narrow and cannot be extended. Furthermore, in assays with e.g., 20-40 fold calibration range, is it necessary to re-adjust positions of QCs? • If in a small study many samples are above the ULOQ then dilution of the samples is appropriate instead of a revalidation of the study. Change of the calibration range should be considered in case of larger studies only. • Line 431: Clarify that this criteria is for small molecule only. • Where wide inter and intra subject variability exists or is observed, for example fast/slow metabolisers it may be necessary at add additional QC's such that subject sub-populations concentrations are better represented.	For LBA this is agreed. The range may not be extended/changed. - The criteria for 2 QCs samples falling within the concentration range is also applicable to LBA. This is not agreed. However, the text has been revised for clarity. The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays. This is correct. Additional QCs may be added. 3 levels is a minimum requirement.
430 - 437	3	Include LBA specific criteria in LBA section. The paragraph is written to make a statement about an	Not applicable to LBA. The text on LBA is revised and a complete new and separate section has been introduced on

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		appropriate use of QC samples - which should cover range of sample concentrations in a study. If this is correct, it should be made more clearly. **432-433.** Regarding statement on assay range being too wide - how is "too wide" defined?	ligand binding assays. Cut-off values are not stated as this will trigger questions why these values were chosen. However the text has been revised for clarity.
430-437	21	When a method is validated the complete range is considered validated and any validated dilution as well. For the majority of LBA calibration range is narrow (20-40 fold) and cannot be extended. **Proposed change (if any):** Delete section 5.3 completely.	Not applicable to LBA. The text on LBA be revised and a complete new and separate section has been introduced on ligand binding assays.
430 - 437	24	Comment: For a study in progress, revalidating the assay over an extended range and "reanalyzing" study samples with this new assay may not be applicable. Investigation of dilution integrity has to be regarded as a more suitable approach (as used in LBAs) to overcome such limitations. Please consider to increase the amount of QC levels in the range of study samples	The guideline does not indicate that the study sample should be reanalysed. 3 concentration levels is a minimum requirement; furthermore the option to include more QCs is in the text (see lines 434-436)
430-437	45	1. Range is validated and dilution as well, this is common practice in ELISA – proposal to keep the validated range and add QC samples if needed. 2. for the majority of LBA calibration range is narrow (20-40 fold) and cannot be extended. Furthermore, in assays 20-40 fold calibration range, is it necessary to re-adjust positions of QCs? 3. Please clarify that these lines apply to small molecule (chromatography) methods only	This is not applicable to LBA. The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
431	23	Large number of samples > ULOQ: not very clear: this would necessitate a better definition of "large number".	Every definition will trigger questions. No definition given. However the text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
431	36	Comment Clarity is sought on the definition of "a large number of samples" in determining whether the assay range should be extended	The text has been revised for clarity.
431 - 432	3	The calibration range is driven by the assay properties, not by the study sample concentrations. If samples are above ULOQ, samples can be diluted. The statement regarding extension of the assay range is very unclear to LBA analysts and should be moved to the LC/MS section.	This is not applicable to LBA. The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
431-432	19	The guideline should be more permissive in this topic. It should be at the discretion of the analysts to decide whether they analyse the samples with dilution or extend the calibration range (and revalidate the method). If dilution is performed in controlled circumstances (e.g. using dilution QC samples), the results can not be questioned.	The text has been revised for clarity. The calibration curve range should be extended, if possible, and QC samples added or their concentrations modified.
431-437	14	Comment: It is unclear what a „large"number and "most of" the analyte concentrations are? This must be better specified. Moreover, the extension of the calibration range is not really possible in case of multi-analyte assays and this should be avoided. It is unclear which limits have to be applied with regard to the upper end of the calibration range. In this context the physico-chemical limits of selected analytical methods have also to be considered. Investigation of dilution integrity has to be regarded as a more suitable approach to overcome such limitations. In general, calibrations ranges should be established in accordance with pharmacokinetic properties of a drug. **Proposed change:** *"Calibration ranges should be established in accordance with the pharmacokinetic properties of a drug, especially*	Cut-off values are not stated as this will trigger questions why these values were chosen. However the text has been revised for clarity. It is not understood why this is not possible? The text has been revised. No criteria have been included on the number of study samples outside the calibration range needing a change in the calibration range.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		for pivotal bioequivalence trials. It should be considered that the concentrations of the unknown samples fall into the validated concentration range. In case concentrations of samples are above the ULOQ and dilution of samples is necessary, dilution integrity of samples has to be proven. If the calibration range is too wide, such that more than two third of the analyte concentrations fall in the lower part of the calibration curve range, an additional QC sample level should be included so that at least 2 QC sample levels fall within the range of concentrations measured in study samples. If the calibration range is extended, the analytical method should be revalidated to ensure accuracy and precision."	
431-437	15	The extension of the calibration range is not really possible in case of multi-analyte assays and this should be avoided.	The comment is unclear.
431-437	22	Comment: Point 5.3 Calibration Range If it is required to extend the calibration range as a result of data we would recommend that it is possible to add an extra QC testing, at the discretion of the analyst, to assure that at least 2 QCs are within the sample concentration range. It should also be allowed that this decision can be taken once the samples of some subjects have been analyzed. The guideline should specify that it is not necessary to re-analyze the subjects analyzed up to that moment.	The guideline does not indicate that the samples should be reanalysed. The possibility of introducing a new QC level is already mentioned in the guideline.
431-437	26	Line 431: Add "The calibration range should be appropriate for the sample concentrations and the project phase of development". Line 437: It is not necessary to re-assay samples just because they have only been analyzed with the original analytical method, and not the updated method with additional QCs or shifted calibration range.	The text has been revised in accordance with the publication *Viswanathan et al., AAPS Journal (2000) 9 (1), art 4.*.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): Modify text as suggested.	
431-437	37	Comment: What does "large number" mean? To facilitate understanding the paragraph should be divided in two separate parts as proposed below. In our opinion, revalidation after narrowing calibration range is not necessary and it should be clearly stated. **Proposed change (if any):** If a large number of the analyte concentrations of the study samples appear to be above the ULOQ (i.e. X% of study samples), extending of calibration curve range should be justified, if possible. If the calibration curve range is extended, the analytical method should be revalidated to define response function and ensure accuracy and precision. ~~The same applies~~ If it appears that the calibration curve range is too wide, such that most analyte concentrations fall in the lower part of the calibration curve range, the ~~calibration curve~~ range should then be narrowed, or an additional QC sample level should be included so that at least 2 QC sample levels fall within the range of concentrations measured in study samples. If the range is narrowed or additional QC level included then the revalidation is not required.	The text has been revised in accordance with the publication *Viswanathan et al., AAPS Journal (2000) 9 (1), art 4*.
432-437	6	Comment: If method is meeting all acceptance criteria in the validated calibration range there is no reason to compromise the method accuracy according to range of values found in study samples. The validated method provides the same validity of results regardless number of samples determined in low or high part of calibration range. **Proposed change (if any):** Delete lines 432-437 from "The same applies "	The text has been revised in accordance with the publication *Viswanathan et al., AAPS Journal (2000) 9 (1), art 4*.
434-435	19	What acceptance criteria should be applied when an	In case an additional QC level is introduced, the same

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		additional QC level is included?	acceptance criteria should be applied as indicated for the other QC samples.
436-437	8	Comment: There is guidance that revalidation is necessary when the curve range is extended. It is also necessary when the range is truncated?	The text has been revised in accordance with the publication *Viswanathan et al., AAPS Journal (2000) 9 (1), art 4*.
436-437	45	Re-validation is needed when the concentration range is extended	The text has been revised indicating that partial validation is needed.
437	47	Comment: Regarding revalidation when a calibration curve range is extended. **Proposed:** Recommend clarification of the necessity to revalidate a method when a partial validation is applicable as written.	The text has been revised indicating that partial validation is needed.
438 - 443	24	Comment: In case of multi analyte assays, providing the number of samples and percentage of total number of samples can be difficult. What is the added value of this requirement?	The difficulty to provide these data is not understood. This is requested for information and clarity during auditing/inspection/assessment.
438-463	2	Comment: • The raw data is the appropriate place to have the re-analysis discussed. The discussion of the re-analyzed samples enlarges unnecessarily the report. • Eliminating the retesting of samples identified as not fitting a pharmacokinetic profile in the absence of guidelines on how to treat such data will introduce some spurious results into the pharmacokinetic analysis and subsequent interpretation. This has the potential to alter study outcomes. This is also inconsistent with international practices. • Pharmacokinetic reason is not acceptable. Just	This is required for the assessment. This is of importance for bioequivalence were manipulation may be of importance. The text has been changed for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		because it looks good does not mean it is correct. Conversely, just because something looks bad does not mean that it is incorrect. Thus, it is appropriate that anomalous results be investigated for systemic errors to confirm or refute the original observation. This is covered in lines 457-459, however, clarity on what is a PK repeat should be detailed. • While the intent to refrain from reanalysis of study samples because of pharmacokinetic reason is understood, in practice it is important to confirm anomalous data; i.e., a sample concentration is inconsistent with the PK profile. Sample tubes may be mislabeled, aliquot errors occur, etc., and confirming the original concentration is the first step in any laboratory investigation. • The reasons for reanalysis should not be this stringent. The term "laboratory investigations" should be replaced by a more specific term, as re-analysis is regularly needed to clarify abnormal data or any analytical doubts. This reanalysis should be ruled in a SOP. • Need to clearly define what is and is not considered to be a PK repeat. **Proposed change (if any):** • Change Line 453 to: Identification of quantifiable analyte levels in pre-dose or placebo samples • Line 454: Statements for internal standard and chromatography are applicable to small molecule only. Clarify this is for small molecule only.	The text has been revised for clarity. It indicates now for bioequivalence studies. It is considered that these are laboratory investigations. Moreover the text has been revised to clarify that this is applicable for bioequivalence studies. This is considered not applicable. This is agreed. The text has been revised accordingly. This is only applicable for chromatographic methods and not for LBA.
438 - 463	3	Include LBA specific criteria in LBA section.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
438-463	26	Line 454: Add run failure because of potentially compromised data quality due to excessive/significant carryover or contamination. Lines 455-459: This section is vague and unhelpful, as written. If the initial analysis of one or more PK samples was shown to be incorrect under appropriate, controlled repeat analysis conditions, a more accurate PK profile would result from the repeat assays. The last sentence should be clarified. Line 463: Reinjection of runs should be allowed if reasonable causes are suspected, such as lack of system equilibration.	In this case the applied method can be questionable. It is agreed that such data may not be used. The text has been rewritten. It indicates now for bioequivalence studies. This is considered covered by line 460.
439	47	Comment: Regarding the statement "Reanalysis of study samples should be predefined in the study protocol or SOP…" should read "The criteria for when reanalysis is acceptable should be defined in the study protocol or SOP." **Proposed:** Recommend that this section be reworded in light of the Comments provided.	The text has been revised and states now 'Possible reasons for reanalysis of study samples and criteria to select the value to be reported should be predefined in the protocol, study plan or SOP..'.
439-441	5	Comment: A percentage total number of samples which could be reanalysed is not given.	This is not possible as the causes for reanalysis may be very different.
439-463	22	Comment: Point 5.4 Reanalysis of study samples Although there are no defined criteria in the guideline for reanalysis due to the IS variability, the EGA is in favour of having those criteria pre-defined in an SOP. Our experience shows that this is considered a crucial parameter for some GCP inspectors, and the final guideline should take this into consideration. The current text considers re-analysis because of	The text of line 439 has been revised.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		pharmacokinetic reason unacceptable. However, re-analysis due to pharmacokinetic data inconsistency may be necessary at times. In the presence of abnormal concentration values which are not consistent with the PK/TK profile it should be allowed to proceed with re-analysis, especially in those studies where estimation of PK parameters is the primary objective (e.g. preclinical PK or human Phase I studies). Re-analysis due to pharmacokinetic reason should only be foreseen under predefined SOP, in a well-documented way and using replicate analysis. We would suggest including a distinction between these situations in the final text of the guideline.	The text has been revised for clarity.
440-443	14	Comment: Reanalysis should not be mixed-up with ISR. From our point of view it is confounding to list initial values in case of an invalid analytical run (e.g. in case of a malfunction of the equipment or wrong spiking of QCs, calibration samples, or internal standard) in the study report. **Proposed change:** Please delete the first 3 bullet points as reasons that reanalysed samples with initial values have to be reported.	This is agreed, Line 441-443 have been deleted and added after line 459. The text has been revised for clarity.
442	8	Comment: The initial value of the repeated sample is asked for, however, in most cases, the original value is unavailable or extrapolated, both of which are not evaluable. **Proposed change (if any):** Remove this statement.	This is agreed, Line 441-443 have been deleted and added after line 459. The text has been revised for clarity.
442-443	47	Comment: Regarding the statements that "the initial value should be provided" and "…justification for the acceptance should be provided." In most of the cases listed in the EMEA guidance document for reanalyzing a sample (lines 444-454), the justification is obvious based upon	This is agreed, line 441-443 have been deleted and added after line 459. The text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		the reason for reanalysis and in most cases, an initial value cannot be obtained. **Proposed:** Recommend adding the words "if possible" prior to the words "the initial value" in line 442 and striking the requirement of providing justification for the reporting of the repeated value in line 443.	
444	47	Comment: The reasons following line 444 are examples, and should read as such, particularly since there are many additional examples **Proposed:** Recommend changing the text to "The following are examples of reasons for study sample reanalysis, but are not limited to:"	The text has been revised in accordance with the proposed change.
444 - 454	24	Comment: **Proposed change (if any):** Please consider to replace "The following reasons can be identified to reanalyse study samples:" by "The following reasons can be identified to reanalyse study samples, e.g.," Please add to the list of reason - any other anomalous results as part of a lab investigation (e.g., sample misidentification). - sample BLQ due to dilution	The text has been revised and states now , ..the following are examples of …'. The example are not limited too.
444-454	45	Remove reasons and refer to SOP	This is not agreed.
444-455	21	While the intent to refrain from reanalysis of study samples because of pharmacokinetic reason is understood, in practice it is important to confirm anomalous data; i.e., a sample concentration is inconsistent with the PK profile. Sample tubes may be mislabeled, aliquot errors occur, etc and confirming the original concentration is the first step in any laboratory	The text has been revised and states now , ..the following are examples of …'. The example are not limited too.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		investigation. **Proposed change (if any):** add italic text to line 453:identification of sample analyte in pre-dose samples or placebo sample *or any other anomalous result,*	
445-446	8	Comment: If the initial values of repeated samples are required, these lines indicate that we would have to report all initial values for failed runs. Since the acceptance criteria has determined that these results are not reliable, what is the value in presenting this data? Also, precision of calibrants are not required as part of batch acceptance.	This is agreed, Line 441-443 have been deleted and added after line 459. The text has been revised for clarity. This is agreed. For calibrator standards precision is not required. The text has been revised.
445 - 454	10	Comment: The differentiation between "improper sample injection or malfunction of equipment" and "poor chromatography" will be rather impossible. In our point of view these are "technical reasons" resulting in reanalysis. **Proposed change (if any):** We would propose the following reasons resulting in reanalysis: Reassays of the samples are carried out in the following cases: 1. The result obtained for the calibration is outside the limits for acceptance. In this case, the whole analytical sequence is rejected and analysed afresh. 2. The result obtained for the QC samples are outside the limits for acceptance. In this case, the following part of the analytical sequence is rejected and analysed anew. Depending on the number of QC samples this could be the whole sequence. 3. The result obtained for an analysed sample is more than 10 % greater than that for the calibrator with the highest concentration of analyte. 4. Samples became mixed up. 5. The chromatograms cannot be evaluated. 6. Technical problems during the analysis. After	It is considered that the current text is sufficient clear.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		solving problems extracts concerned will be injected one further time.	
447	23	What are the criteria to be used regarding Internal Standard (IS) variability? What about stable labelled Internal Standards?	Line 439 has been revised. This issue is considered covered in general and in line 447/448.
447-448	14	Comment: From our point of view it is unclear what "significantly different" means? If this requirement for the internal standard would be taken into consideration, then within every run a statistical significance test on the signal of internal standard has to be performed. A difference in the internal standard response does not necessarily mean that the study sample results are invalid. As earlier discussed in the section on matrix effects, the isotope-labelled internal standard compensates for matrix effects in most cases and this should have been investigated during method validation. Therefore in our opinion, sound scientific criteria for the internal standard response evaluation cannot be defined.	As indicated, these criteria have to be predefined by the laboratory in a SOP.
447-448	23	Is this criteria of re-analysis considered to be critical point by point aspect of the validation? e.g. acceptance limits?	As indicated, these criteria have to be predefined by the laboratory in a SOP. It depends on the criteria set in the SOP.
447-448	47	Comment: Regarding the inclusion of an internal standard response assessment. If an internal standard response is significantly different from the internal standard response of the calibration standards and/or QC samples, then the method is likely not robust, and probably should be investigated. "Significantly different" is a subjective evaluation rather than an objective evaluation. **Proposed:** We recommend that rather than require an investigation for "significant differences" in IS response, that the guidance require that the degree of acceptable variation	As indicated, these criteria have to be predefined by the laboratory in a SOP. It depends on the criteria set in the SOP.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		in this response be described in the applicable method and if there are variations in the IS response, an investigation into the cause should still be considered.	
450-452	37	Comment: See comment to limes 407-411 **Proposed change (if any):** ~~the obtained concentration is above the ULOQ or below the run's LLOQ, in runs where the lowest standard sample has been rejected from a calibration curve, resulting in a higher LLOQ compared with other run's,~~	The proposed changes are considered not acceptable.
450-452	47	Comment: Regarding reasons for reanalysis of samples. Samples can also fall below the LLOQ and need to be repeated if they were diluted samples that happened to be overly diluted. **Proposed:** Recommend including this scenario as a possible reason for reanalysis for study samples.	This issue is considered resolved by adding a new introduction sentence. The text has been revised and states now , ..the following are examples of …'. The example are not limited too.
453	37	Comment: additional word **Proposed change (if any):** identification of ~~sample~~ analyte in pre-dose samples or placebo sample,	The text has been revised for clarity.
453	49	A repeat is only required if the analyte identified in the pre-dose sample is above LOQ. All other samples below LOQ are reported as BLOQ.	The text has been revised for clarity.
454	9	**Proposed change (if any):** Please add an additional bullet point relating to ligand binding assays: "exceeding the % CV criterion of duplicate study sample measurement for large molecule, if the concentration falls within the validated assay range"	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
454	21	Propose to add "Poor replicate analysis for ligand binding assays"	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
454	23	**Proposed Change:** Propose to add "Poor replicate analysis for ligand binding assays".	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
454	25	Comment: The term "Poor chromatography" is too unspecific. **Proposed change (if any):** The method should describe the definition of what is "good chromatography", for example by comparing to a reference chromatogram.	Line 439 has been revised indicating that the reason for reanalysis should be predefined in a protocol, study plan or SOP.
454	47	Comment: Regarding the reason of "poor chromatography" as a possible reason for reanalysis. Poor chromatography is a subjective evaluation (rather than an objective evaluation). **Proposed:** Recommend this reason be removed from the guidance or a more definitive discussion of what constitutes "poor chromatography" be added to the guidance.	Line 439 has been revised indicating that the reason for reanalysis should be predefined in a protocol, study plan or SOP.
455	16	Comment: Eliminating the retesting of samples identified as not fitting a pharmacokinetic profile in the absence of guidelines on how to treat such data will introduce some spurious results into the pharmacokinetic analysis and subsequent interpretation. This has the potential to alter study outcomes. This is also inconsistent with international practices **Proposed change (if any):** Please provide guidelines to the pharmacokineticist in the recommended approach for dealing with concentration data that is inconsistent with the pharmacokinetic profile.	The text has been revised for clarity. It indicates now for bioequivalence studies.
455	3	Statement does not apply to real world situations. Should state that sample reanalysis for PK reasons is acceptable provided there is clear documentation of the reason, and that the documented reason and request	The text has been revised for clarity. It indicates now for bioequivalence studies.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		come from the pharmacokineticist. A priori criteria for reanalysis of samples due to PK repeats, or failed CV etc. should be pre-defined, perhaps in an SOP.	
455	14	Comment: In practice it is important to confirm anomalous PK data; e.g. sample concentration is inconsistent with the PK profile: sample tubes may be mislabelled, aliquot errors occur, etc… **Proposed change:** Please delete the sentence. Alternatively, this sentence could be limited to bioequivalence studies.	The text has been revised for clarity. It indicates now for bioequivalence studies.
455	23	An outlier /aberrant value can be observed on the pharmacokinetic profile and be the reflection of a handling error during the analytical phase. If an aberrant value is detected on a pharmacokinetic profile and if we do not reanalyse, it means considering that an aberrant value is acceptable. Isn't it critical? We could demonstrate a bioequivalence where there isn't and vice versa. It appears strange that it is not allowed to re-analyse samples that are considered by the responsible scientist to be out of the expected PK/TK profile? There may be situations when a sponsor of a contracted study would specifically request this in order to verify accuracy of the 1st result obtained. If such procedures for repeat analysis are documented in approved SOPs, is it really a problem? The SOP (and GLP) would require a pre-defined, standard approach, justified in the raw data and report. It is perhaps naive to indicate "Normally the reanalysis of study samples because of a pharmacokinetic reason is not acceptable." There will always be questions coming from our TK/PK colleagues about specific sample data. The new EMEA Guidance should state that it is permissible for BA to re-run a sample for TK/PK reasons provided that there is a clear reason for it provided by the pharmacokinetic colleagues and that it is	The text has been revised for clarity. It indicates now for bioequivalence studies.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		documented prior to start of the re-assay. If maintained, The text should clarify if the reanalysis might be included in the final results or definitively not.	
455	36	Comment; "Normally re-analysis of study samples because of a pharmacokinetic reason is not acceptable" It is suggested that the guidance is modified to align with FDA CDER BMV guidance and other best practice docuemnts which indicate that analysis is permitted provided that the reasons and a PK/TK driven rationale is made prior to analysis and data acceptance and reporting. If certain study types preclude this (such as Bioequivalence studies)s then this should be emphasized and detailed in the guidance..	The text has been revised for clarity. It indicates now for bioequivalence studies.
455	39	This requirement assumes that the operator makes no errors, but this is simply not true. The scientific approach is to reanalyse all aberrant values, otherwise a single error could change the outcome of the entire study. Moreover, this requirement is not consistent with that on l. 453, because the reanalysis of the pre-dose sample is also a reanalysis because of a pharmacokinetic reason. **Proposed change (if any):** Replace with the sentence "The samples highly aberrant from the typical pharmacokinetic profile should be reanalysed".	The text has been revised for clarity. It indicates now for bioequivalence studies.
455	40	Comment: In my opinion, PK re-assay is very important, especially for BE studies. I have managed a bio-analytical laboratory for over 15 years. From my experience, there are cases where an abnormal can occur and reasons cannot be found. I have seen values 10 times higher than the highest Cmax or highest standard. Should these values be included in the BE assessment, they would jeopardize the outcome,	The text has been revised for clarity. It indicates now for bioequivalence studies.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		especially the Cmax. That means a good product may be judged a failed one, or vice versa. I am aware of the concern of the agency and fully support your concern, but I believe PK re-assay should be done. The key is that it has to be done properly to avoid any data mishandling. **Proposed change (if any):** Should PK re-assay be deemed necessary, it must be selected and re-analyzed before the statistical analysis. The re-analysis would be done in duplicate in order to confirm the concentration of the re-assay sample.	
455	49	This line however specifically states that reanalysis is not permitted for pharmacokinetic reasons, but from the above it should be obvious that the "kinetic" deviation might be due to haemolysis. Reanalysis should be permitted to confirm that the value is in actual fact correct and the possibility of not reporting this value should be considered. Reanalysis merely to prevent recurrence of similar problems in the future (457 – 459) is of no use to the current study being analysed and a decision regarding the abnormal value needs to be taken.	The text has been revised for clarity. It indicates now for bioequivalence studies.
455-456	7	Comment: Re-analysis because of pharmacokinetic reason should be possible provided the request is raised by the pharmacokineticist and not the bioanalyst. The request is to be documented. It is acknowledged that the necessity for PK repeats in BE studies is lower than for early Phase studies because the PK profile is already known	The text has been revised for clarity. It indicates now for bioequivalence studies.
455-457	37	Comment: The word normally is not clear. If there are specific cases where reanalysis for PK reasons is acceptable it should be clearly stated. In our opinion the lack of PK based reanalysis can strongly influence quality of data and bias outcome of BE studies (especially on clearly outlying point with significantly higher concentration than expected from PK	The text has been revised for clarity. It indicates now for bioequivalence studies.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		profile, which may become outlying Cmax). The reason for analysis may not be clear during sample preparation or analysis, but the final concentration may strongly suggest some kind of error. As statistically 95% of results should be within mean +/- 2 SD, we suggest deviation over 30% of expected value as acceptance criteria for PK-based sample re-analysis. **Proposed change (if any):** Reanalysis of study samples because of a pharmacokinetic reason must be justified and clearly described for each sample (e.g. single sample deviation from pharmacokinetic profile of more than 30% of expected value, inconsistent data for elimination rate constant calculation etc.).	
455-459	6	Comment: This requirement assumes that the operator makes no errors, but this is simply not true. The scientific approach is to reanalyse all aberrant values, otherwise a single error could change the outcome of the entire study. Moreover, this requirement is not consistent with that on l. 453, because the reanalysis of the pre-dose sample is also a reanalysis for pharmacokinetic reason. The FDA guideline also allows reanalysis because of inconsistent pharmacokinetic data "Reasons for repeat analyses could include repeat analysis of clinical or preclinical samples for regulatory purposes, inconsistent replicate analysis, samples outside of the assay range, sample processing errors, equipment failure, poor chromatography, and **inconsistent pharmacokinetic data**. Reassays should be done in triplicate if sample volume allows." **Proposed change (if any):** Replace with the sentence "The samples highly aberrant from the typical pharmacokinetic profile should be reanalysed. Reassays should be done in triplicate if sample volume allows."	The text has been revised for clarity. It indicates now for bioequivalence studies.
455-459	21	It is naïve to state that reanalysis of study samples	The text has been revised for clarity. It indicates now for

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		because of a pharmacokinetic reason is not acceptable. There will always be questions coming from our TK/PK colleagues about specific sample data. This Guideline should state that it is allowed for BA to re-run a sample for TK/PK reasons provided that there is a clear reason for it which is documented prior to start of the re-assay (to be driven by PK/TK group, not by BA). **Proposed change (if any):** Replace line 455-459 by "Normally reanalysis of study samples from bioequivalence studies because of a pharmacokinetic reason is not acceptable as this may affect and bias the outcome of such a study. In general reanalysis of study samples because of a pharmacokinetic reason may be allowed only when there is a clearly documented reason prior to reanalysis. Additionally reanalysis might be considered as part of laboratory investigations, to identify possible reasons for results considered as abnormal and to prevent the recurrence of similar problems in the future."	bioequivalence studies.
455-459	21	Reanalysis can be conducted as part of a laboratory investigation. **Proposed Change:** Clarify what values should be reported if samples are reanalyzed as part of an investigation.	The text has been revised for clarity. It indicates now for bioequivalence studies.
455-459	19	The abnormal concentration values which are not consistent with the PK/TK profile should be allowed to be reanalysed, especially in those studies where estimation of PK parameters is the primary objective (e.g. preclinical PK or human Phase I studies). Reanalysis due to pharmacokinetic reason should be performed only under predefined SOP, in a well-documented way and using replicate analysis.	The text has been revised for clarity. It indicates now for bioequivalence studies.
455-459	31	Comment: Repeats for pharmacokinetic reasons are generally not accepted according to guideline. However, pharmacokinetic repeats are allowed as part of	The text has been revised for clarity. It indicates now for bioequivalence studies.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		laboratory investigations. In case, the repeated analysis performed as part laboratory investigations show another more plausible value, it is unclear how this situation should be handled. Should the implausible value nevertheless be reported in the study report? **Proposed change:** It is suggested to allow repeats for analytical reasons under the following prerequisites : - It has been defined in the study protocol a priory. - Samples to be repeated for pharmacokinetic reasons are identified by a blinded pharmacokineticist. - Such samples are identified as defined in the study protocol, analytical protocol are SOP a priori.	
455-459	32	Where physiologically impossible data are generated then an SOP defining gross anomalies should be developed eg predose samples with high levels of drug: an intermediate sample of a profile with zero levels but bracketed by samples with high levels of drug.	The text has been revised for clarity. It indicates now for bioequivalence studies.
455-459	45	1. Lines 455-459 discourage reanalysis of samples because of PK reasons. PK sometimes is able to point out to laboratory errors, e.g. switching of samples and laboratory investigation take place. While lab investigation are allowed here, it is not clear what to do when the results are confirmed that errors took place. We should keep a process that allows unbiased use of the correct results, rather than using the knowingly wrong results 2. are "pharmacokinetic anomalies" absolutely excluded as reasons for re-assays? 3. We suggest that the following statement may be too broad. "Normally reanalysis of study samples	The text has been revised for clarity. It indicates now for bioequivalence studies.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		because of a pharmacokinetic reason is not acceptable." 4. There should be some consideration for investigation, and if it appears that a systematic error (e.g. sample handling error) was made, a reanalysis can occur. Or if there is a statistically based decision that a result is outside the "PK" profile, then it should be allowed. Also, there is no mention of reassays due to a known technical error. We suggest that reanalysis is acceptable if a known technical error occurred and is adequately documented	
457-459	8	Comment: It is appreciated that it has been taken into account in the guidance that reanalysis might need to occur as part of laboratory investigations.	The comment is noted.
457-459	37	Comment: this not a requirement and should be deleted. **Proposed change (if any):** ~~However reanalysis might be considered as part of laboratory investigations, to identify possible reasons for results considered as abnormal and to prevent the recurrence of similar problems in the future.~~	This is considered a recommendation. The proposed change is not agreed.
457-459	49	If reanalysis is performed as part of a laboratory investigation and it is proven that there was an error, what result is then reported?	The text has been revised for clarity. It indicates now for bioequivalence studies.
460-463	50	Comment: Re-injection of the samples can be made in case of instrument failure, if re-injection reproducibility and on-injector stability have been demonstrated during validation. An experimental design for the conduct of re-injection reproducibility and on-injector stability should be incorporated in this section.	The experimental design is considered an issue to be covered by an SOP and not by this guideline.
464	23	It is already necessary to have an SOP for integration of chromatograms.	Yes; it is considered that is has been made clear how integration issues are handled.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
464 - 466	3	Not appropriate for LBA. Place in LC-MS section.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding assays.
465	23	There may be more sources of bioanalytical data than chromatograms**.** It should also be possible to reflect the chromatogram integration method in the study report **Proposed change (if any):** Review the wordings so the text can accommodate the different possibilities.	This is a specific concern regarding chromatograms. The study report is prepared after completion of the study. Therefore it should be included in a SOP.
465-466	14	Comment: From our point of view it is unclear what the generic content of such a SOP on chromatogram integration could be? It seems more appropriate to write a brief section on integration of chromatograms in the method description. **Proposed change:** Propose to rephrase to: *"The procedure with re-integration of chromatograms should be described in a SOP."*	The text has been revised and 're-integration' has been added.
465-466	15	Unclear description for re-integration. **Proposed change:** The procedures performing automatic or manual re-integration of chromatographic signals should be described in a SOP.	That is what is meant.
465-466	50	Comment: Chromatogram integration should be described in a SOP. Any deviation from this SOP should be discussed in the analytical report. Consistency in applicability of validated integration parameters should be discussed in details in this section, and deviation from validated integration parameters should be justified and backed up with the experimental	It is considered that this should be included in the SOP.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		data.	
465-466 519 539	47	Comment: Regarding the requirement to include deviations from SOPs and/or methods in the report. **Proposed:** Recommend that these requirements be clarified from a GLP perspective, particularly since only those types of deviations that impact the quality and integrity of a GLP study need to be included in the report.	It is considered that the text regarding this issue does not need revision. However, the revised text is more elaborate for clarity, with reference to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010)
467	23	If we consider that the long-term stability in the storage conditions is validated, what is the scientific rationale to redo these analyses? What would be the percentage of samples to be reanalysed?	Long term stability data is a different subject in validation than ISR.
467	50	Lines 467; Comment: Incurred samples re-analysis Based on analyte characteristics and bioanalytical requirements, incurred samples re-analysis may be a part of method validation, if samples are available to prove the reproducibility of method. Additional details on minimum no. of samples and Acceptance criteria and ISR failure will be useful for users.	The guideline indicates the acceptance criteria.
467-472	23	How many samples should be reanalysed in percentage? It is important to specify this since in pre-clinical studies this could possibly impact the number of animals to be used. An indication of number of study samples to be re-analyzed would be useful.	No absolute value is given, however as a guide the guideline indicates now: 10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples.
467 - 477	24	Comment: Chapter should be adapted according to EBF recommendations. as described in "Incurred Sample Reproducibility: Views and Recommendations by the European Bioanalysis Forum, P. Timmerman et al" (PDF attached)	It is considered that the revised paragraph covers sufficiently the recommendations of the EBF. Regarding the criteria, 20% difference from the mean would be in line with the White paper (Fast et all., AAPS Journal, DOI

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		In summary, EBF recommends the performance of ISR for the following studies: • In support of method validation:: • First time into new matrix animal or human • First time into new target population: First Patient study • Disease state changes in patient population • Upon any major method change (as per SOP) • First time use of an existing method in a new laboratory • Whenever scientific reasons require retesting of ISR (e.g. special populations based on scientific rationale) • As part of regular process check: • All BE studies • Incidental check in any other studies (human PK or DDI study) EBF recommends following acceptance criteria: • 2/3 of repeat values within 80-120 % of original value (chromatographic assays) • 2/3 of repeat values within 70-130 % of original value (ligand binding assays) • EBF recommends number of reanalyzed samples to be at the discretion on the scientist and based on scientific judgement	10.128/s12248-009-9100-9)
467-486	2	Comment: • It is unrealistic to expect that samples from every subject in a clinical study will tested for ISR. • For human study samples, ISR for every subject on study? Could this text be clarified? • ISR has been well discussed and presented in the AAPS white paper as 10- 5% depending on study size small to large. This should be accommodated in this Guideline also.	As a guide the guideline indicates now: '10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.'

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		• To perform ISR for every subject is stretching beyond reason. Performing ISR for first-in-human (FIH) and BE studies would sufficiently demonstrate assay reproducibility and anything beyond that should be issue driven.	
467 - 486	3	Include LBA specific criteria in LBA section. Greater guidance (re: expectations for ISR) should be spelled out. Section should be re-written.	The text on LBA is revised and a complete new and separate section has been introduced on ligand binding
467-486	6	Comment: Health Canada discontinued this practice in 2003. FDA discussed this issue at AAPS Workshop and published the results in The AAPS Journal in 2007. However, no outcome of this discussion was implemented to FDA guideline. Besides that the current paragraph does not define: the total number of samples to be re-assayed; the action to be taken toward the studies in which the method was used when re-analysis does not meet the acceptance criteria; what time period should be between the assay for the study and re-assay. Such analytical investigation should be required for PK Phase I studies with new molecules but the strict requirement for BE studies is not based on scientific ground. **Proposed change (if any):** Delete this paragraph	The issue of ISR is introduced again and therefore included in this guideline.
467-486	12	Comment: The document do not clearly mention about the reanalysis of the incurred samples from the preclinical toxicokinetic studies involving small rodents such as rat. The sample volume in these studies is typically very low (2 aliquots of 100 µl each). In instances where the second set of aliquot is used for repeat analysis for any rejected batches for analytical related issues or samples lost during shipment from multisite studies, the incurred sample reanalysis will be	Pooling is not acceptable. It is clear that no repeats can be carried out because of insufficient sample. However, this is not applicable to all animals.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		not possible to perform for individual samples. **Proposed change (if any):** The recommendation is to pool all the samples from a dose group and sex to get single sample per dose and sex that can be used as incurred sample in two different occasions to estimate the reproducibility of the results with defined acceptance criteria	
467-486	18	Comment: Health Canada discontinued this practice in 2003. FDA discussed this issue at AAPS Workshop and published the results in The AAPS Journal in 2007. However, no outcome of this discussion was implemented to FDA guideline. Besides that the current paragraph does not define: the total number of samples to be re-assayed; the action to be taken toward the studies in which the method was used when re-analysis does not meet the acceptance criteria; what time period should be between the assay for the study and re-assay. Such analytical investigation should be required for PK Phase I studies with new molecules but the strict requirement for BE studies is not based on scientific ground. **Proposed change (if any):** Delete this paragraph	The issue of ISR is introduced again and therefore included in this guideline.
467-486	26	Line 470: Not clear how concomitant medications affect assay reproducibility. Line 472: "The extent of testing depends on the analyte and the study samples….." Can this be clarified? Line 472: What is meant by "over a certain time period"? Line 475: Change the word "accurate" to "reproducible". This would then be consistent with the definition on Line 567.	Line 470: Concomitant medications can affect the matrix effect. Line 472: As an example, if it is known that back-conversion takes place, this will be a critical issue for ISR. Line 472: The text has been revised and indicates now 'in separate runs at different days…'. Line 475: The text has been revised and this line is deleted. accordingly. Line 483: As a guide the guideline indicates now:

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Line 483: "from several subjects" Can some additional guidance be provided on the approximate percentage of samples expected to be re-analysed?	'10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.'
467-486 Incurred Sample Reanalysis	36	Comments Calcification is sought for this whole section. In particular that the requirements differing between clinical and preclinical studies are given. Additional clarity on expectations for performing ISR including providing recommendation on the number of samples per study that should be selected for ISR is sought. The expected methodology of determining the % variability (per the ISR publication) should be given and the number of samples assessed for ISR. Guidance on how diluted samples are handled is requested. It is suggested that this section is consistent with the Workshop Report and Follow-Up—AAPS Workshop on Current Topics in GLP Bioanalysis: Assay Reproducibility for Incurred Samples—Implications of Crystal City Recommendations (published in AAPS Journal April 2009 10.1208 s12248-009-9100-9) and other publication s on this topic.	The text has been revised. As a guide the guideline indicates now: '10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.'
467-486	37	Comment: It is definitely true that QC and calibration samples do not mimic actual study samples. In our opinion incurrent sample reanalysis does not provide additional information on results quality. Moreover, accuracy of incurred samples reanalysis is by definition impossible because the true value is unknown. If bioanalytical method is biased with constant error, the incurred samples reanalysis does not show it. If significant back-conversion takes place there will be significant trends in reported data and an additional experiment is unnecessary. In our opinion incurred samples reanalysis does not provide additional information on protein binding nor influence of concomitant medications, while homogeneity of sample	The results obtained should be reproducible. As data have been shown that reproducibility/accuracy may be a problem in studies due to the issue mentioned in line 469, the need for ISR is to confirm the reproducibility of the study outcome. Therefore ISR is needed.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		should be ensured by application of proper procedures of sample handling. In case of bioequivalence studies, which are definitely comparative studies, the request for the analysis of incurred samples in our opinion inconsistent with aim of study: method should allow comparison of two data sets (i.e. bioanalytical results) not determination of absolute concentration. The constant ratio of Cmax between test and reference drug is of primary interest rather than identical concentration measured during 2 separate analysis of each sample. The recommendation to study incurred samples seems inconsistent with not allowing sample re-analysis due to PK reasons. We suggest to delete this or change acceptance criteria for BE studies (criteria should be based on constant T/R concentration ratio rather than constant concentration).	
467-486	39	Is there any scientific evidence that the incurred sample reanalysis could reveal problems with different protein binding, back-conversion of known and unknown metabolites, sample inhomogeneity or concomitant medications or is this just speculation? These problems can occur during sample processing, but hardly during storage. Thus, the reanalysis of the samples would hardly detect these problems. Moreover, it is not clear why in bioequivalence (BE) studies the reanalysis of incurred samples should always be carried out. BE studies represent relative measurement (Test/Reference); the possibly increased inaccuracy has no effect on the outcome of BE study at all. If precision would be worsened, then the chance of concluding bioequivalence is lowered, which mean decreased risk for a patient. After reanalyses two values will be available for some samples - how will they be treated for statistical evaluation of the BE study? And finally, the requirement for the difference of two measurements to be within 20% is more strict than that	Yes, the scientific data have been published. Especially for BE studies the accuracy of the concentration is of importance as the study is used to make a conclusion. If for instance back conversion of metabolite occurs, it may influence the intra subject CV and thereby the outcome of the study. The first value should be taken into account for evaluation. The second one is only for purpose of ISR. There was an international consensus on this criteria (see

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		for QC samples. Incurred sample reanalysis can be performed only with real patients samples. In BE studies a serious problem arise: either a pilot study should be performed or study samples from the respective BE study can be used. The first approach will bring ethic problems, the second is impossible when validation must be finished before the analysis of study samples. **Proposed change (if any):** Delete the entire section.	White Paper). ISR can be an addition to validation.
467-486	41	Comment: No percentage is given for ISR. **Proposed change (if any):** Should include % e.g. >1000 samples – 10%; 1001-2000 samples – 5-10%; >2000 samples – 5%. It should be better defined for which studies ISR should be done, and for which it is recommended.	As a guide the guideline indicates now: '10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.'
467-486	47	Comment: The incurred sample reanalysis (ISR) section does not address the following: • Number of samples to be analyzed and the respective conditions • Reporting • Use of data obtained for ISR during the reanalysis of a run using a multi-analyte method **Proposed:** Recommend that this section be reworded in light of the Comments provided.	- As a guide the guideline indicates now: '10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.' - This is indicated in line 542. - This can be used, however this is only for one analyte and for one subject.
467-489	49	More detailed information is required for analysis of incurred samples. What samples must be used? How many samples should be analysed? If, as stated, it is required to use samples at maximum and in the elimination phase, then this can only be done after a number of volunteers have been analysed so that the	- As a guide the guideline indicates now: '10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		maximum concentration samples can be identified. If a Cmax value is a suspected "abnormal" value, then what sample must be used? If this method is used then the number of incurred samples to be analysed for a large study (eg: 48 volunteers, 4 phases) is abnormally high	phase.' - It is agreed that PK should be known. - The first run is for evaluation purpose, the second one for ISR. If a sample is suspected to be abnormal, a repeat analysis can be carried out as indicated in section 5.4.
468-486	19	The number of samples to be reanalysed should be defined in the guideline. In case of incurred sample reanalysis what values should be accepted for PK/TK analysis when the second analysis confirms the first result and when the second result differs significantly from the first result (it can be occurred in 33% of the repeats by the guideline)?	As a guide the guideline indicates now: '10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.' The first analysis prevails. If large differences are observed this should start investigations even if the overall acceptance criteria are met. The text has been revised regarding this aspect.
468-486	22	Comment: Point 6 Incurred Samples reanalysis We would suggest that the number of samples to be reported should be defined, as well as which samples should be reanalyzed. We propose that a minimum number of 50 samples, or 5% of the study samples corresponding to the Cmax and the elimination phase, whichever is greater. The ISR test should not be regulated in pilot studies, although it can be a good opportunity to carry out the test prior to the pivotal study.	As a guide the guideline indicates now: '10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.' The guideline does not request this.
471-472	21	**Comments**: It is not clear what is meant by "over a certain time period". To avoid issues related to stability, it is recommended to evaluate incurred sample reproducibility as close as possible to the first analysis,	The text has been revised and indicates now 'in separate runs at different days…'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		in a separate batch and not over a certain period. **Proposed change (if any):** To replace "It is therefore recommended to evaluate accuracy of incurred samples by reanalysis of study samples over a certain time period" by "It is therefore recommended to evaluate accuracy of incurred sample through reanalysis as close as possible to the first analysis, in a separate batch".	
471-472	23	It is unclear what you mean by "over a certain time period". To avoid issues related to stability, it is recommended to evaluate incurred sample reproductiibility as closest as possible from the first analysis, in a separate batch and not over a certain time period.	The text has been revised and indicates now 'in separate runs at different days…'.
471-472	40	Comment: The requirement for ISR is not clear. Please define required sample size for ISR. **Proposed change (if any):** 30% of total subjects or a maximum of 10 subjects should be selected for ISR.	As a guide the guideline indicates now: '10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.'
472ff	17	Comment: "Reanalysis of samples over a certain time period" → there is no guidance in the document whether samples to be reanalyzed within studies have to be covered by stability testing.	The text has been revised and indicates now 'in separate runs at different days…'.
472-474	31	Comment: The guideline states that the extent of testing incurred samples by reanalysis (ISR) depends on the analyte and the study samples, and should be based upon in- depth understanding of the analytical method and analyte. Information about the number of samples to be analysed is not provided. **Proposed change:** It is suggested to specify the number of samples to be	As a guide the guideline indicates now: '10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.'

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		analysed in ISR, e.g. for bioequivalence studies 5% of samples, but not more than 100 samples per study.	
472-477	45	No useful comment (over a certain time period – remove this part)	The text has been revised and indicates now 'in separate runs at different days…'.
475-476	12	Comment: how much deviation dictates an investigation?	There is no consensus about this. Therefore no guidance can be given.
475-477	8	Comment: It is unclear if an investigation would be required for individual ISR sample failures, or overall failure.	This should be done in both cases.
475-477	47	Comment: Regarding investigations of incurred sample reanalysis (ISR) results. It is unclear whether the expectation of an investigation refers to an overall ISR failure (e.g., more than 67% of ISR do not meet the defined acceptance criteria) or individual ISR failures (one or more samples but less than 67% do not meet the acceptance criteria). **Proposed:** Recommend that this section be reworded in light of the Comments provided.	This should be done in both cases.
478	7	Comment: If pharmacokinetic parameters represent the primary endpoints of a study ...Please delete this sentence (as other regulatory authorities did) or consider re-phrasing such as "***If pharmacokinetic parameters represent the primary endpoints of a study, incurred sample analysis may be considered***" There may be several studies in healthy volunteers with PK parameters as primary endpoint. There is no need to perform ISR in each of these studies. **Proposed change (if any):** Please delete the sentence in line 478-479 or consider re-phrasing such as "***If pharmacokinetic parameters represent the primary endpoints of a study,***	It is considered that the current revised text is sufficient clear.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		incurred sample analysis may be considered"	
478 - 481	3	For ISR studies, include first use of assay and platform change.	In case a platform has changed a partial partial validation should be carried out. Regarding first use: it should be done as early as possible.
478-482	14	Comment: From our point of view it is rather important to investigate ISR in studies with different subject populations than in every study with pharmacokinetics as primary endpoint. The purpose of ISR is to assess robustness of a method. Therefore, a more consequent comparative evaluation of bioanalytical and pharmacokinetic results in a few selected studies seems more promising to mirror robustness of a method instead of perpetually repeating ISR. For animal studies we recommended to investigate ISR in non-GLP pharmacokinetic studies as they are earlier available than toxicokinetic studies. Furthermore, those studies are conducted at lower doses and should be therefore more sensitive with regard to instability of unknown metabolites and the relative impact on signal of parent compound. Saturation of P450 isozymes in high dose (toxicokinetic) studies has also to be taken into consideration resulting in a lower quotient of metabolite to parent compound. Please refer also to EBF publication on ISR. **Proposed change:** Propose to rephrase to: "*ISR should be investigated in studies with different subject populations as far as possible (e.g. samples of subjects from highest dose group of first-into-human study, samples of first patient trial, samples of subjects from special populations, every time scientific reasons recommend evaluating ISR). The robustness of an analytical method should optionally also be investigated in samples of subjects from drug-drug-interaction studies unless implausible*	The text has been revised according to the proposed changed for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		pharmacokinetic results have been obtained. For bioequivalence studies analysis of incurred samples should always be carried out. Samples should be taken from every subject or patient unless otherwise justified to avoid bias in selection of the samples. For samples of animal pharmacokinetic studies it is sufficient to investigate ISR in the main toxicology species (e.g. rat and dog or monkey).”	
478-484	46	Comment: It is not clear whether incurred sample analysis should be performed for every subject in bioequivalence studies. In lines 479-481 it is stated that “For (human) study samples evaluation of incurred samples should be carried out for every subject or patient population, unless otherwise justified”, whereas in lines 482-484 it is reported that “It is recommended that study samples are obtained from several subjects close to the expected maximal concentration and in the elimination phase”. We would like EMA to specify the percent of subjects whose sample will be analysed as incurred samples. **Proposed change (if any):** For (human) study samples evaluation of incurred samples should be carried out for 20% of subjects or patient population, unless otherwise justified.	The text has been revised for clarity. As a guide the guideline indicates now: ‘10% of the samples should be reanalysed in case the number of samples is less than 1000 samples and 5% of the number of samples exceeding 1000 samples. Furthermore, it is advised to obtain samples around Cmax and in the elimination phase.’
478 - 486	24	Comment: Please consider to update " In toxico-kinetic studies it is sufficient to address this issue once per species" to " A proper evaluation of incurred sample reproducibility and accuracy needs to be performed on each species used for Good Laboratory Practice (GLP) toxicology experiments. " (harmonization with the FDA white paper)	The text has been revised for clarity.
479	47	Comment: Regarding the statement “In toxico-kinetic studies it is sufficient to address this issue once per species.” **Proposed:**	The text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Recommend revising this to include "once per species per matrix" to be consistent with the Crystal City meeting consensus.	
479-481	23	We recommend to adopt the same approach as the FDA and to do it in a subset of samples. **Proposed Change:** To replace "For (human)v study samples evaluation of incurred samples should be carried out for every subject or patient population, unless otherwise justified." by "For (human)v study samples evaluation of incurred samples should be carried out on a subset of samples".	The text has been revised for clarity.
479-481	45	We don't agree to this requirement: "should be carried out for every subject". Might be excessive to evaluate ISR samples for every subject and could lead to unblinding a study? Support to perform ISR on a per patient population; however suggest calling it disease indication **Proposed change (if any):** For (human) study samples evaluation of incurred samples should be carried out in each disease indication, unless otherwise specified	The text has been revised for clarity.
479-482	21	To perform ISR for every subject is stretching beyond reason. Performing ISR for FIH and BE studies would sufficiently demonstrate assay reproducibility and anything beyond that should be issue driven. **Proposed change (if any):** For (human) study samples evaluation of incurred samples should be carried out ~~for every subject or patient population, unless otherwise justified~~ *at least for first-in-human and bioequivalence studies.* ~~For bioequivalence studies analysis of incurred samples should always be carried out.~~ *Incurred sample analysis beyond these studies may be considered and should be issue driven.*	The text has been revised for clarity.
480	3	Clarify that ISR will not be carried out on EVERY subject,	The text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		but on each new population.	
480	47	Comment: Regarding the statement "…incurred samples should be carried out for every subject or patient population…" **Proposed:** Recommend that the term subject be clarified as this is not consistent with the Crystal City meeting consensus.	The text has been revised for clarity.
480 and 483	11	Comment: "every subject" and "several subjects" are at conflict.	The text has been revised for clarity.
481-482	31	Comment: In the guideline, it is stated that incurred sample reanalysis (ISR) is always required for bioequivalence studies. **Proposed change:** In case several bioequivalence studies are required for the same product using the same bioanalytical method with the same range and in the same laboratory, it is suggested that ISR only needs to be performed once for the entire study package.	This is agreed. But study samples should be obtained from the different studies and identified as such.
481-482	46	Comment: What does it mean that "incurred sample analysis should be evaluated as early as soon as possible"? **Proposed change (if any):** Analysis of incurred samples should be performed only for the samples from the first analytical runs.	It meant that ISR evaluation should not be done after the study is completed, but investigate ISR early in the study phase.
482	21	Incurred sample analysis should be evaluated as early as possible. There is a lack of homogeneity with line 472 "over a certain time period"	The text has been revised for clarity. .
485 - 486	3	Math behind calculation for the % difference from the mean should be clarified: [repeat - original]/mean	There are no maths behind this; it is in line with the White paper consensus.
487	8	Comment: The entire section is unclear about what pertains to the analytical report and what pertains to the	The text has been revised. A clear separation between

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		validation report. **Proposed change (if any):** The section should be subdivided more clearly.	validation report and analytical report is included.
487	11	Comment: The whole section is a little bit messy. **Proposed change (if any):** Please consider separate it into two subsections, such as validation and study sample analysis	The text has been revised. A clear separation between validation report and analytical report is included..
487	21	Section 7 is entitled "Study Report", but contains references to "final report", "validation report", "study report", and "analytical report", which makes difficult to interpret the reporting requirements. **Proposed change (if any):** We would recommend to modify the text to clearly define both the common and separate requirements for the Validation Report and sample analysis Study Report and to modify the Section title to be clearly reflect the content of the section.	The text has been revised for consistency. Furthermore, the text has been revised and a clear separation between validation report and analytical report is included.
487	23	"7. Study Report" – Section is entitled "Study Report", but contains references to "final report", "validation report", "study report", and "analytical report", which makes it difficult to interpret the Validation Report and sample analysis Study Report requirements. **Proposed change:** To modify text to clearly define both the common and separate requirements for the Validation Report and sample analysis Study Report; to modify Section title to be representative of both the Validation Report and sample analysis Study Report.	The text has been revised for consistency. Furthermore, the text has been revised and a clear separation between validation report and analytical report is included.
487 - 490	24	Comment: It is not clear whether criteria are described for a study report, a bioanalytical report or a method validation report. Independent criteria would be defined for validation and study report. We would propose to generate a table summarizing which topics are needed in which report and the same for raw data (please harmonize with FDA white paper	The text has been revised for consistency. Furthermore, the text has been revised and a clear separation between validation report and analytical report is included. Study director is the GLP term, however this has been deleted

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		There is likely to be confusion around the use of the term Study Director. Study Director would be fine for an assay validation study if ever claimed GLP.	in the revised text.
487-544	2	Comment: • This section deals with reporting and it is not always clearly delineated between what statements relate to validation reports and what is for study sample analysis reports. • Line 491: What is meant "complete documentation of the protocol"? Only relevant parts of the protocol should be reflected in the study report. • SOPs should absolutely not be appended to the report. They would unnecessarily enlarge the report. SOPs can always be submitted to the Health Authority whenever needed and at any time. • It is sufficient to have the sample tracking in the raw data. There is no added value to have this type of information in the study report. It's not the meaning of a study report to reflect as much as possible the raw data. In case of questions insight into the raw data is warranted at any time. • With the exception of BE & BA studies where 5-20% chromatograms are appended, only representative chromatograms should be appended to a report. • As the Agency acknowledges that sample analysis is not necessarily conducted according to GLP principles and as a bioanalytical report is considered a component report, it is not always the study director signing bioanalysis reports.	The text has been revised and a clear separation between validation report and analytical report is included. The relevant parts should be included and parts can be referred. However this also refers to conduct and evaluation. The text has been revised to prevent misunderstanding. Nowadays studies are generally submitted in eTCD format. The report should clearly indicate how the study samples were handled and stored. Sampling tracking in general should be described and not per subject; the latter can be obtained from the raw data. For BE 20% is agreed. For other studies relevant chromatograms can be appended and should be available on request. The text has been revised for clarity. The term 'study director' has been deleted in the revised text. Regarding this issue, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) This is acknowledged.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		This should be reflected in the text. • Performing bioanalytical method validation applying GLP standards is not consistent with OECD-GLP and FDA standards. Moreover, GLP requires independent Quality Assurance auditing of the study; which again is inconsistent with FDA requirements. • Not all raw data is electronically available. • Line 494: There is no previous definition of an SOP. In this context, it is interpreted to mean the approved bioanalytical method. • Line 505: Inclusion of the preparation dates for all solutions and QCs is not needed if the duration of storage for stability testing is presented. • Line 521: Inclusion of all individual result concentrations for all experiments, especially when presenting the results of triplicate measurements for stability at two or more concentrations, would not necessarily add any more value than presenting the mean and standard deviation or percent CV. • Lines 532-533 "table of all analytical runs with analysis dates and results, indicating which samples have been analyzed in which analytical run" is absolutely excessive and has no added value. • Line 519: What about minor SOPs (not the method SOP) deviations that do not have an impact and can be addressed in the raw data? For example, pipette use and calibration SOP.	The term SOP reflects the standard operating procedure. Definition will be included. Preparation dates should be included to verify the storage time period. This is not agreed. The individual data gives direct information as is the case for all other data. During evaluation/assessment the reader can directly interpret the data. This is not the case with only presenting mean / CV data. This is considered not excessive. It should be clear from the report in which run which sample has been included. This is acknowledged, but the question is always what a minor deviation is. The text has been revised. Regarding this issue, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010) The text has been revised for clarity. The text has been revised accordingly.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): • Change Line 503 to: "Reference standards (Origin, Batch number, Certificate of Analysis, stability and storage conditions)" • Replace lines 488-490 by: "In case sample analysis is performed in accordance with GLP principles either the study director or principal investigator or contributing scientist should sign and date the final report to indicate acceptance of responsibility for the validity of the data and to indicate the extent to which the study complies with the principles of good laboratory practice." • Replace lines 491-493 by: "The validation report should include complete documentation of the protocol or reference to the SOP followed, conduct and evaluation of the analysis." • Replace line 495 by: "All source data should be available in its original format and available upon request."	
487 - 544	3	Change section title to **Validation Report**, and only discuss contents of validation report.	The title has been changed, and the text has been revised. A clear separation between validation report and analytical report is included.
487-544	13	Comment: The very detailed description of the features of the study report use the terminology of GLP studies. It seems questionable if the responsible scientist, who developed and validated the bioanalytical method and analyzes the clinical samples, can take over the responsibility of a GLP study director, especially when samples derived from clinical studies are concerned. Here, rather the data will be reported back to the data management center. The guideline also requires that individual data should be provided in electronic format, which is not further specified. Further definitions, especially for	The text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		individual data should be given in the Definitions section <545 ff.>	
487-544	14	Comment: From our point of view the whole chapter should be better organized, e.g. 7.1 Validation report 7.2 Study report In the present version it is a mix of both. A more clear distinction should be made between the data that should be reported and the data that should be included to raw data. There is likely to be confusion on the term „Study Director": in non-clinical and clinical studies the Study Director is the „in-life"contact person and the bioanalyst is the „bioanalytical responsible person".	The text has been revised and a clear separation between validation report and analytical report is included. The guidance included is in accordance with other publications.
487-544	15	The whole chapter should be rearranged, e.g. 7.1 Validation report 7.2 Study report	The text has been revised and a clear separation between validation report and analytical report is included.
487 to 544, Section 7	16	Comment: This section deals with reporting and it is not always clearly delineated between what statements relate to validation reports and what is for study sample analysis reports. **Proposed change (if any):** Please clearly separate the discussion on validation and study sample analysis reports.	The text has been revised and a clear separation between validation report and analytical report is included.
487-544	21	It would be advisable to try and harmonize reporting requirements with other Agencies and in particular the FDA to avoid the need to produce multiple reports.	The EMA considers harmonisation an important issue. Harmonisation issues have been taken into account as much as possible.
487-544	26	Line 494: Relevant SOPs should not need to be appended; cross-reference should suffice. Line 505: Preparation dates of QCs not necessary for reports. Line 510: Calibration results only necessary in report for	Line 494: The text has been revised for clarity. Line 505: Preparations dates should be included to verify the storage period of the study samples. Line 510: This is agreed, but should be available.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		passing runs. Line 513: QC results only necessary in report for passing runs. Line 517: Dilution results only necessary in report if dilutions evaluated. Line 532: Not always practical in later stage clinical studies to have table of analytical runs indicating samples assayed. Line 544: "All chromatograms of …". Difficult to include every chromatogram from a subject specified prior to analysis when some samples from the same subject are later assayed in different runs due to: 1) dilution for initial value above ULOQ, 2) requested for PK repeat (re-assayed later in duplicate), and 3) analytical reason.	Line 513: This is agreed, but should be available. Line 517: This is agreed, and it also accounts for matrix effect. Line 532: Comment noted. Line 544: This is agreed. The text has been revised for clarity.
487-544	45	1. Please replace this chapter by what is actually in the FDA guidance. Please align to avoid need to issue 2 different study reports, one for FDA and one for EMEA 2. Separate between validation and study report and add in case one more chapter.	1. The EMA considers harmonisation an important issue. Harmonisation issues have been taken into account as much as possible. 2. The text has been revised and a clear separation between validation report and analytical report is included.
487-544	52	Comments: The very detailed description of the features of the study report uses the terminology of GLP studies. It seems questionable if the responsible scientist, who developed and validated the bioanalytical method and analyzes the clinical samples, can take over the responsibility of a GLP study director, especially when samples derived from clinical studies are concerned. Here, rather the data will be reported back to the data management center. The guideline also requires that individual data should be provided in electronic format, which is not further specified. Further definitions, especially for individual data should be given in the Definitions section <545 ff.>.	The text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
488	36	Comments The text references the sign/dating of study report by a Study Director and compliance to the GLP rules, As this is only applicable to preclinical sample analysis clarification in the text is recommended	The text has been revised and the term study director has been deleted. Furthermore, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010)
488-492	47	Comment: Regarding the reference to the "study director", the requirement to report "the extent to which the study complies with the principles of good laboratory practices", "…validation report should include complete documentation of the protocol…", and "in accordance with the principles of GLP-rules…". **Proposed:** Similar to earlier comments for lines 58 and 69-70, these statements do not provide alternative guidance for reporting for those facilities that cannot or do not adhere to the principles of GLP when conducting bioanalytical method validation studies. Recommend rewording the section to include, for example a "responsible scientist" as an alternative to a study director and documentation of a validation SOP as an alternative to the protocol as well indicating the GLP compliance statement is only required if the facility complies with GLPs.	The text has been revised and the term study director has been deleted. Furthermore, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010)
488	45	Study director should sign the study report, is very specific to non-clinical study reports per FDA GLP regulations (Code of Federal Regulations 21 Part 58.185-14b) – and suggest that this be stated.	The text has been revised and the term study director has been deleted. Furthermore, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010)
488-544 Study report	36	Comments It is suggested that the introduction of this section should include a summary of what documents are expected to be provided and in what formats as a means of providing clear guidance on the expectations of the agency.	The text has been revised and considered sufficient clear now.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
488-490	21	As the Agency acknowledges that sample analysis is not necessarily conducted according to GLP principles and as a bioanalytical report is considered a component report it is not always the study director signing bioanalysis reports. This should be reflected in the text. **Proposed change (if any):** Replace lines 488-490 by: In case sample analysis is performed in accordance with GLP principles either the study director or principal investigator or contributing scientist should sign and date the final report to indicate acceptance of responsibility for the validity of the data and to indicate the extent to which the study complies with the principles of good laboratory practice.	The text has been revised and the term study director has been deleted. Furthermore, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010)
488-490	23	**Comments**: Bioanalysis in a GLP study: The Study Report is commonly referred to as the report which the Study Director is responsible for. In a multi-site study there is a Principal Investigator (PI) who is responsible for the bio-analytical part of the study. The PI provides the report for the bioanalytical part and this report is amended to the Study Report. **Proposed change (if any):** Study Reports are a central element in the OECD GLP therefore consider review and update of this part so it fits the roles and responsibilities concerning Test Facility/Site management, Study Director/Principal Investigator and QA Program.	The text has been revised and the term study director has been deleted. Furthermore, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010)
488, 492	8	Comment: All terminology and reference to GLP is difficult to apply, since method validation does not fall under the scope of the GLPs.	This is agreed.
488-496	19	There are several phrases in the text for the report, such as "final report", "validation report", "report", "study report" and "analytical report", that causes confusion. It should be clarified which sentences apply to which report. **Proposed change (if any):** We suggest to use "pre-	The terminology will be brought in line. Furthermore, the text has been revised and a clear separation between validation report and analytical report is included.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		study validation report" for reporting validation results and "bioanalytical study report" for reporting trial sample analysis results.	
488-496	22	Comment: Point 7 Study Report There are several phrases in the text for the report, such as "final report", "validation report", "report", "study report" and "analytical report", that causes confusion. It should be clarified which sentences apply to which report. **Proposed change (if any):** We suggest to use "pre-study validation report" for reporting validation results and "bioanalytical study report" for reporting trial sample analysis results.	The terminology will be brought in line. Furthermore, the text has been revised and a clear separation between validation report and analytical report is included.
488-544	22	Comment: Point 7 Study Report (relevant SOPs) The SOP of the analytical method should not be added, nor other SOPs relevant in the analytical report, since they are a part of the company's know-how. These SOPs should be available upon request of the agencies at times of inspections.	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
489	4	Delete 'is'	The text has been revised.
489	11	Comment: delete "is"	The text has been revised.
489	23	Delete "is".	The text has been revised.
489	31	Typing error: …"the extent to which the study ~~**is**~~ complies with…	The text has been revised.
489	37	Comment: additional word **Proposed change (if any):** the study ~~is~~ complies with	The text has been revised.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
491	6	Comment: **Proposed change (if any):** Replace the "validation report" with "study report"	The text has been revised for clarity and consistency.
491	9	Clarify if line 491 requires that audits be conducted for validation studies as this is not a requirement under current regulations. **Proposed change (if any):** For clarity, suggest providing separate sub-sections for validation and study reports.	The text has been deleted. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010).
491	18	Comment: **Proposed change (if any):** Replace the "validation report" with "study report"	The text has been revised for clarity and consistency.
491-493	11	Comment: It is not clear. Include protocol per se in the report?	The text has been revised for clarity and consistency.
491-493	14	Comment: It is not clear what "complete documentation of the protocol"is? Please avoid any reference to GLP unless GLP compliance is required.	The text has been deleted. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010).
491-493	21	Performing bioanalytical method validation applying GLP standards is not consistent with OECD-GLP and FDA standards. Moreover, GLP requires independent Quality Assurance auditing of the study; which again is inconsistent with FDA requirements. **Proposed change (if any):** Replace lines 491-493 by: "*The validation report should include complete documentation of the protocol or reference to the SOP followed, conduct and evaluation of the analysis*."	The text has been deleted. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010).
491 - 493	24	Comment: **Proposed change (if any):** We propose to rephrase	The text has been deleted. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples."

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		the sentence to: "The validation report should include sufficiently precise description of the bioanalytical method, validation experiments as they were designed, acceptance criteria and a discussion of results generated".	(EMA/INS/GCP/532137/2010).
491-493	45	1. GLP only for tox studies, not for method validation. 2. No protocol for validation, should be done according to SOP or protocol	The text has been deleted. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010).
491-544	15	The bioanalytical information specified in the protocol/study plan should be mentioned within the bioanalytical report. A study report cannot replace study audits and it should be avoided to add all relevant SOPs the report, because almost all SOPs may be relevant. SOPs are internal documents and should not be distributed with any report.	The text has been deleted. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010).
492	23	Should the validation be in compliance with GLP and ICH guidelines? The GLPs are designed to be applicable to non-clinical safety studies and validation has always been excluded since this occurs independently of the study. No other method validation is considered necessary to be performed under GLPS e.g. all medical laboratory analyses. Some authorities are specifically requiring that routine analysis of clinical samples must be performed according to GCP.	The text has been deleted. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010).
492	37	Comment: additional word **Proposed change (if any):** with the principles of GLP~~–rules~~	The text has been deleted. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010).
493	21	"… conducted audits/inspection should be included in the report." – it is not clear if this reference is solely for sample analysis Study Reports or also for Validation Reports. When full GLP compliance is not required for method validations, then validation specific	The text has been deleted. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010).

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		audit/inspections are not required either. **Proposed change (if any):** We would suggest to revise the sentence to read as follows: "..conducted audits/inspection should be included in the study report for GLP compliant study sample analysis ."	
493	23	**Comments**: ".. conducted audits/inspection should be included in the report." – it is not clear if this reference is solely for sample analysis Study Reports or also for Validation Reports; If full GLP compliance is not required of method validations, then validation specific audit/inspections are not required. If QA statements are however required, OECD GLP prescribes who is responsible for the reports and how to perform and report audits/inspections in a QA Statement. **Proposed change (if any):** Consider review and update of this part so it fits the roles and responsibilities concerning the QA Program. Modify the text to "..conducted audits/inspection should be included in the report for GLP compliant study sample analysis."	The text has been deleted. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010).
494	4	Comment: Do not agree with SOPs being in a study report as appendices but happy for copies to be in the study file. In addition not sure which SOPs are being requested. Is it the SOP on repeat analysis, SOP on reintegration, preparation of the calibration standards and QCs, operation of the instrument etc.? SOPs are controlled confidential documents specific to an organisation. Once they appear in a report it can be difficult to control their circulation and confidentiality. Proposal: List the SOPs that you expect to be incorporated in the study file supporting a study report.	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	6	Comment: Such SOP includes the intellectual property of lab/CRO and as such cannot be disclosed in the report. **Proposed change (if any):**	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Delete this line	
494	7	Comment: **Proposed change (if any):** Please delete the sentence or re-phrase "***SOP's for relevant analysis specific procedures may be appended to the study report***""	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	9	Please clarify this sentence. If this means to append the specific SOP or Analytical Procedure for the method (i.e. one document describing the method), this is appropriate. However, this sentence could also be interpreted as requiring all SOPs used during the analysis to be appended, which would often require most of a labs SOPs to be included. We believe this would be unnecessary, non value-added and inefficient.	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	14	Comment: A study report cannot replace study audits and it should be avoided to add all relevant SOPs to the report, because almost all SOPs may be relevant. Since in general there are many SOPs on analytical procedures within a pharmaceutical company this would result in an inadequate „blow-up"of every single study report. Moreover, SOPs are internal documents and describe in addition to scientific issues also management procedures of a company. Therefore, they should be only available on a special request of an authority and not be distributed with any report. **Proposed change:** Propose to rephrase to: *"SOP's for relevant analysis specific procedures should be appended to the study report if <u>referenced</u> in the study report."*	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	16	Comment: There is no previous definition of an SOP. In this context, it is interpreted to mean the approved bioanalytical method. **Proposed change (if any):** Please define SOP or	A definition of SOP will be included in the guideline.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		differentiate the use from SOPs related to routine laboratory activities (e.g., use of a balance).	
494	18	Comment: Such SOP includes the intellectual property of lab/CRO and as such cannot be disclosed in the report. **Proposed change (if any):** Delete this line	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	19	Adding the relevant SOP's to the study report can be problematic since they are written in local languages in most of the cases. Additionally, it is not well-defined what are the "SOP's for relevant analysis specific procedures". The SOPs should be available upon request of the agencies.	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	21	We would suggest that SOPs should be readily available on site for inspection if required and that only a detailed description of the methods is included in the report.	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	23	"SOP's for relevant analysis specific procedures should be appended to the study report" - SOPs may not be available when validation studies start. Any SOPs available are already summarized in the content of the report. They are available in the documentary system of the company. If we start to append the SOPs in the report with this guideline, the drift could be to append the analytical SOPs in the study reports for other types of analyses, and eventually for all activities performed within a study. This requirement should be looked at again since it may be in conflict with the GLPs. Why should SOP's for relevant analysis specific procedures be appended to the study report when SOPs are written in the national language and not in English? It does not appear acceptable to attach SOPs in our reports since we may have many documents and from country to country, the SOPs, written in local languages, would necessitate translations into English. This does not seem realistic.	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed Change: To replace "SOPs for relevant analytical specific procedures should be appended to the study report" by "The report includes the exact description of assay used in the study".	
494	24	Comment: We believe that adding and referencing SOPs as validation report appendix has no added value.	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	32	We agree with the rationale but the report should use other processes to ensure transparency ie all processes should be summarised eg as per study plan (appended to report) and the SOP referenced as appropriate otherwise the report will become VERY large	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	36	Comment Clarity is requested on the text referencing the need for SOP's for relevant analysis specific procedures to be appended to study reports. The text in line 499 may be relevant as this refers to the applied analytical method	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	37	Comment: In our opinion the clear description of relevant analysis specific procedures should be placed in the report. It may be written in the main text of the validation report as well. Moreover, there may be a problem when SOP is in other language than typical for study report, i.e. the report is written in English, while SOP in Polish. In this case translation of whole SOP seem to be necessary and SOPs in both languages (?) have to be attached. The strict requirement for SOP attachment seems to be an overregulation. **Proposed change (if any):** The relevant analysis specific procedures should be described in the study report or appendices. or ~~SOP's for relevant analysis specific procedures should be~~	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		~~appended to the study report.~~	
494	39	SOP's for relevant analysis specific procedures represent the intellectual property of the analytical laboratory and should not be disclosed in the study report. **Proposed change (if any):** Delete this sentence	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	41	Comment: Relevant SOP's should be appended to the report. This is not possible to do, as SOP's are intellectual property of a company. Furthermore, it may result in appending >100 SOP's to each report, since all SOP's are relevant. **Proposed change (if any):** Omit the requirement to append the SOP's.	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	45	1. No SOPs should be provided. They are available on-site and archived in the GLP-archive (available for inspection) 2. Line 494 requires inclusion of SOPs in study reports. It is not clear which: all SOPs should be included and if they required to be repeated in submission of multiple reports or only once and then referred to in subsequent report	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.
494	48	Comment: What does the guidance mean by : " relevant analytical specific procedures" ? Does it mean : Only the detail of the method or well the SOP relating to the way to conduct a validation study , the acceptance criteria applied for the validation study (even if this acceptance criteria are already described in the study plan)? **Proposed change (if any):** Please detail which are expected to be described in these SOP (that have to be appended to the validation	The text has been revised for clarity. It states now that depending on the detailed information in the report, reference can be made to the SOPs.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		study report).	
495	14	Comment: This should be restricted to concentration result data and not be required for raw data. Readability of electronic data long term may well be an issue, which is why some companies claim their „raw data" paper or a PDF rendition of the paper.	The text has been revised and indicates now that all source data should be availble on request.
495	21	Not all raw data is electronically available and it would be unrealistic to include such requirement. **Proposed change (if any):** We would recommend to replace line 495 by: *"All source data should be available in its original format and available upon request."*	The text has been revised and indicates now that all source data should be availble on request.
495	23	"All individual data should be available in electronic format to be provided upon request". What would be the electronic format, and what would the support need to be: CD-Rom or access in the software for the chromatographic interpretation? In any case, This requirement should be deleted since there is no GLP requirement to have individual data in electronic format (i.e. for paper raw data).	eTCD format is the principle submission format for regulatory dossiers. In the section 3 Legal basis, reference is made to the reflection paper for laboratories that perform the analysis or evaluation of clinical trial samples." (EMA/INS/GCP/532137/2010).
495	24	Comment: **Proposed change (if any):** Since electronic raw data are not a requirement, we would appreciate the following language for this sentence: "All individual data should be available to be provided upon request." or this deletion.	The text has been revised and indicates now that all source data should be availble on request.
495	34	Many CRO consider that the first print of the results are the "raw data". **Proposed change:** Please Remove "in electronic format" in the text	The text has been revised and indicates now that all source data should be availble on request.
495	40	Comment: Please clarify what was meant by: "all individual data…" . Do you mean chromatography, batch acquisition data or the report? My interpretation	The text has been revised and indicates now all source data should be availble on request.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		is that the chromatography and batch acquisition data require special software to read.	
495	45	It is nice to see in line 495 to see that requirement for availability of electronic data is clearly defined. However, paper is still to be accepted as raw data	The text has been revised and indicates now all source data should be availble on request.
496	9	Reference to analytical protocol in this sentence is potentially confusing. For validations, deviations to the validation protocol should be recorded. For sample analysis, deviations from the study protocol, analytical procedure, or SOPs should be discussed in the analytical report.	The terminology will be brought in line. Furthermore, the text has been revised and a clear separation between validation report and analytical report is included.
496	45	Does this refer to Bioanalytical reports?	The terminology will be brought in line. Furthermore, the text has been revised and a clear separation between validation report and analytical report is included.
497 - 507	24	Comment: Information on calibration standards and QC samples (preparation, preparation dates and storage locations) should only be available in the raw data.	This information should be included to verify preparation and the storage of these samples.
498	11	Comment: Does it mean "summary of validation results"?	The text has been revised and a clear separation between validation report and analytical report is included.
498	21	We would suggest that a summary of validation performances instead of the validation report be presented in a table format. **Proposed change (if any):** We recommend to replace "Summary of the validation report in a table format" by "summary of validation performances."	The text has been revised. Presentation of the summary in table format is not specific indicated anymore.
498	23	We suggest a summary of the validation performances instead of validation report in a table format.	The text has been revised and it states now 'summary of the validation performances'.
498	45	Mandatory table format for summary of validation report?	The text has been revised and it states now 'summary of the validation performances'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
499	37	Comment: summary or details? **Proposed change (if any):** ~~summary~~ details on the applied method	The text has been revised and it states now 'details of…'.
501	11	Comment: "summary" and "details" are kind of contradictory.	The text has been revised and it states now 'details of…'.
501	23	"Summary of the assay procedure" The purpose of validation study is to determine the procedures of measurement. **Proposed change:** "details of the assay procedure.	The text has been revised and it states now 'details of…'.
501-502	37	Comment: summary or details? **Proposed change (if any):** details on the assay procedure (analyte, IS, ~~details of the~~ sample pre-treatment, extraction and analysis),	The text has been revised and it states now 'details of…'.
503	21	We would suggest that if the batch, CoA and stability information is on file and traceable through the report, this detail does not be added to the report	This is considered relevant information for the reviewer.
503	45	CoA to raw data, not into the report, If the batch, CoA and stability information is on file and traceable through the report, suggest that this detail not be added to the report	This is considered relevant information for the reviewer.
505	16	Comment: Inclusion of the preparation dates for all solutions and QCs is not needed if the duration of storage stability testing is presented. **Proposed change (if any):** Please remove the requirement for inclusion of preparation dates.	This is considered relevant information for the reviewer. This information should be included to verify preparation and the storage.
506	36	Comment. Addition of text or tables in the validation report documenting failed validation runs (identity, date, reason for failure) to the validation report is suggested.	This is indicated by the guideline in line 518.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
507	7	Comment: For method validation sample tracking is not as important as for study sample analysis **Proposed change (if any):** Please delete statement	This is agreed. The text has been revised and sample tracking is deleted.
507	14	**Proposed change:** Please delete the statement because for method validation sample tracking is not as important as for study sample analysis.	This is agreed. The text has been revised and sample tracking is deleted.
507	21	We would suggest to delete this statement because for method validation sample tracking is not as important as for study sample analysis. We would suggest to include sample preparation recommendation instead of sample tracking (conditions and duration of storage).	This is agreed. The text has been revised and sample tracking is deleted.
507	23	We suggest sample preparation recommendations instead of sample tracking (conditions and duration of storage).	The text has been revised and sample tracking is deleted.
507	32	In general the term sample tracking is inappropriate for validation samples. An alternative would be "it should be clear how the validation samples (QCs) were prepared, handled and stored prior to analysis".	This is agreed. The text has been revised and sample tracking is deleted.
507	37	Comment: This is described in lines 504-505 (during validation only QC and calibration samples are used). **Proposed change (if any):** ~~sample tracking (conditions and duration of storage),~~	This is agreed. The text has been revised and sample tracking is deleted.
507	40	Comment: Please clarify the requirement for sample tracking. Would a statement of the sample storage condition and duration be sufficient for the report?	The text has been revised and sample tracking is deleted.
507	45	QC sample tracking for validation is not needed	This is agreed. The text has been revised and sample tracking is deleted.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
507	48	Comment: What does the guidance mean by: " sample tracking" for the validation study report. Does it mean that - the short term or the freeze/thaw stability table should be presented with the date and the time of the QC preparation; of the QC storage (in each condition) and of the QC analysis – - the date/ time of QC preparation and of QC analysis should also be reported for the QC used for the within and between run accuracy and precision?	The text has been revised and sample tracking is deleted.
508 - 520	24	Comment: Needs a clear separation here between validation reports, study reports and documentation in study file. We would propose to generate a table summarizing which topics are needed in which report and the same for raw data (please harmonize with FDA white paper). Details of sample tracking should be available in study file. "table of all analytical runs with analysis dates and results, indicating which samples have been analyzed in which analytical run" would not be a requirement. For most studies typical chromatograms are sufficient (e.g.: double blank, blank, LLOQ, low and high QC, low and high standard, Cmax study sample and Cmin study sample). It should be clarified for which study types a larger number of chromatograms is requested (BE). In validation reports, chromatograms of selectivity samples should be presented.	The text has been revised and a clear separation between validation report and analytical report is included. Harmonisation issues have been taken into account as much as possible. The text on sample tracking has been revised and sample tracking is deleted. The text regarding 'table of all analytical runs…' has been revised for clarity. The request for chromatograms has been adapted: for bioequivalence studies 20% is still requested, however for other studies relevant chromatograms can be appended and should be available on request.
510 - 512	35	Comment: The values outside acceptance criteria should be identified in the table **Proposed change (if any):** Add at the end of the sentence "values outside acceptance criteria should be clearly marked"	The text has been revised according to the proposed change.
510-512	37	Comment: This point is not clear. In our opinion calibration range is set during validation and does not	The response function is not a static issue, but differs per

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		change between batches (see comments to lines 407-411). The response function should be set during evaluation of at least 3 calibration curves and should not be changed during further experiments (so it should not differ between curves). The within- and between-run accuracy and precision are evaluated on QC samples not calibration curves as described in guideline in lines 182-223. **Proposed change (if any):** table of calibration results of all analytical runs, including ~~calibration range, response function,~~ back-calculated concentrations, ~~within- and between-run precision and accuracy~~,	calibration line. For calibration standards precision is not applicable. The text has been revised and precision has been deleted.
510-512	39	Within- and between-run precision and accuracy cannot be determined on calibration samples. **Proposed change (if any):** Delete the words "within- and between-run precision and accuracy".	For calibration standards precision is not applicable. The text has been revised and precision has been deleted.
511	6	Comment: Within-run precision cannot be determined on calibration samples **Proposed change (if any):** Delete "within- and"	For calibration standards precision is not applicable. The text has been revised and precision has been deleted.
511	18	Comment: Within-run precision cannot be determined on calibration samples **Proposed change (if any):** Delete "within- and"	For calibration standards precision is not applicable. The text has been revised and precision has been deleted.
511	23	For the calibration curve there should not be "within-run" precision?	For calibration standards precision is not applicable. The text has been revised and precision has been deleted.
513 – 514	35	Comment: The values outside acceptance criteria should be identified in the table	The text has been revised according to the proposed change.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): Add at the end of the sentence "values outside acceptance criteria should be clearly marked"	
513-514	37	Comment: This point is not clear. There is no requirement in guideline to use extra QC samples for batch acceptance during method validation, so which QC should be tabled? Or is it only table for accuracy and precision experiments? **Proposed change (if any):** ~~table of QC results of all analytical runs (within- and between-run precision and accuracy),~~	The guideline indicates QC results, meaning also for the extra QC samples, if appliable.
517	23	For the "Matrix effect" complete with "if applicable" as from this text it applies to the mass spectrometric analytical methods.	The text has been revised and 'if applicable' is added.
518	37	Comment: This point is not clear. What is definition of "deviating result"? Is it an outlier? We suggest that validation report should inform if acceptance criteria were passed or not with 0/1 kind of answer. There is no need to explain each outlying result – this is statistically proven that the presence of outliers is something natural. **Proposed change (if any):** ~~deviating results obtained during validation with full justification of the action taken,~~	The text has been revised for clarity.
518	40	Comment: There are times that deviating values are due to pure human error and do not necessarily require justification. In my opinion, if 90% of QC samples are within 15%, the deviating results are not from the method. It is most likely due to human sample process error. No justification should be required. **Proposed change (if any):** If more than 10% of QC samples from a validation have deviating results, a full justification would be required.	10% deviating results due to human error would be a critical issue. The text has been revised for clarity.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
518	45	**Proposed change (if any):** Suggest changing the word deviating to "unexpected results obtained…."	The text has been revised according to the proposed change.
520-522	23	"All measurements with the individual calculated values have to be presented" . Do we also have to present all rejected data and if yes, is it necessary to include them or not in the statistical calculations, knowing that this will obviously produce a bias in the results? Although the study sample report is separate from the validation report it is often not a stand-alone entity. It is usually a sub-report to an overall study report. For samples analyzed under GLPs, the report would be generated and the compliance statement signed by a Principal Investigator. It would include a QA statement and be forwarded to the Study Director for incorporation into the main report as an Appendix. **Proposed Change:** The situation for phase reports in pre-clinical studies should be addressed.	As indicated in line 518, deviating results should be reported. The data to be reported depends on the rejected data. The text has been revised and a clear separation between validation report and analytical report is included.
521	16	Comment: Inclusion of all individual result concentrations for all experiments, especially when presenting the results of triplicate measurements for stability at two or more concentrations, would not necessarily add any more value than presenting the mean and standard deviation or percent CV **Proposed change (if any):** Please permit reporting of stability results as the mean and an appropriate measure of variability.	Individula data should be presented to verify the data.
521	36	Comments: Clarification is requested regarding statement: All measurements with the individual calculated concentrations have to be presented in the report Does this mean raw data has to be presented? Above statement (line 495) suggests that individual data should be available and provided upon request. These	The guideline indicates that all obtained concentration data have to be reported. Not the raw data.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		two statements should be aligned.	
522-523	21	Although the study sample report is separate from the validation report it often is not a stand-alone entity. It is usually a sub-report of an overall study report. For samples analyzed under GLPs, the report would be generated and the compliance statement signed by a Principal Investigator. It would include a QA statement and be forwarded to the Study Director for incorporation into the main report as an Appendix. **Proposed change (if any):** We recommend to delete "as a separate report" .	The text has been revised and a clear separation between validation report and analytical report is included.
522-523	37	Comment: To improve the understanding of guideline "separate report" should be named, for example as bioanalytical report. We suggest that this separate report will be an appendix to the main study report (e.g. CTD 5.3.1.2. Comparative and Bioequivalence Study Reports) as it is directly associated with the study, not with validation (CTD 5.3.1.4. Reports of Bioanalytical and Analytical Methods for Human Studies). **Proposed change (if any):** Furthermore a specific detailed description of the analysis of the study samples should be written as a separate bioanalytical report and should include at least… The report may be prepared as separate document or an attachment to PK report	The text has been revised and a clear separation between validation report and analytical report is included.
522-523	45	Is an analysis report (bioanalytical report) needed for all studies? Suggest that it be required for Bioequivalence studies or studies where PK is a primary endpoint and then suggested (but not mandatory) for all other studies.	The text has been revised and a clear separation between validation report and analytical report is included.
526	11	Comment: It is not clear by "method reference". Does it mean "literature reference" or "reference materials"?	This means literature or the validation report, if applicable. The text has been revised for clarity.
528	23	Are the details needed in a report for sample tracking : dates of receipt and content, storage location? This	These data should be presented to verify the data by

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		information is necessary but should only need to be available for on site inspections since it will not be interpretable without information concerning the freezer calibrations etc.	reviewer.
529	9	**Proposed change (if any):** Recommend that only an overview of sample tracking is included in the report.	These data should be presented to verify the data by reviewer.
531	11	Comment: Does it mean the contents of an analytical run?	Contents of an anaytical run is meant. The text has been revised for clarity.
531	23	Can you clarify "set-up of the analytical run". Is it the list of runs or the typical content of a run?	Contents of an anaytical run is meant. The text has been revised for clarity.
532	23	Table of all analytical runs with analysis dates and results indicating which samples have been analysed in which analytical run. Is this level of details needed?	It is considered that this is not detailed.
532-533	21	It is not always feasible to include a table indicating which samples were analyzed in a particular run. This is especially difficult in clinical studies where samples are analyzed in batches containing multiple subjects at different points of protocol. It is not clear what the value of this inclusion is since this information is retained in raw data of study analysis.	This is not understood. It should be known in which run which study samples have been included.
532-533	37	Comment: The sentence is not clear. What does "analytical run result" mean? **Proposed change (if any):** table of all analytical runs with analysis dates and general identification of studied samples (i.e. samples from subject No. X).	Contents of an anaytical run is meant. The text has been revised for clarity.
532-533	45	It is not always feasible to include a table indicating which samples were analyzed in a particular run. This is especially difficult in clinical studies where samples are analyzed in batches containing multiple subjects at different points of protocol. It is not clear what the value of this inclusion is since this information is retained in raw data of study analysis.	The text has been revised and states now 'table identifying all analytical runs and study samples with run dates and results'.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
541	45	Suggest that a reassay table only include those samples that were requested to be reassayed. For samples that were reassayed due to analytical reasons (e.g failed run), this should not be required.	This is agreed. The text has been revised to clarify this.
541	48	Comment: In case of re-assay due to samples > ULOQ, is it necessary to mention the original value taken into account that they could only be presented for informative purpose since outside of the validated range) – In this case, the mention ">ULOQ" is it suitable? Is it possible to mention only the re-assay value? (i.e. the final value with the diluted sample) **Proposed change (if any):** "re-assay (table of sample identification , reason for re-assay, original (only if the first value obtained is within the validated range (or ILQ) and if the corresponding assay series belongs to a validated one) and re-assay values)	This is for informative purpose. The text has been revised for clarity.
542	11	Comment: Does it mean incurred sample re-analysis or reproducibility?	ISR is meant. The text has been revised and indicates now 'incurred sample reanalysis'.
543	9	Please clarify what is meant by the range of 5% to 20%? We support including 20% of chromatograms for pivotal bioequivalence studies, but suggest that this requirement is reconsidered for other types of study. **Proposed change (if any):** We suggest to either remove the requirement, or indicate "example" chromatograms should be included. Please also add a comment that this does not apply to ligand binding assays.	The text has been revised and indicates now that for bioequivalence studies 20% of the chromatograms should be included and for other studies the relevant chromatograms can be appended and should be available on request. The text on LBA be revised and a complete new and separate section has been introduced on ligand binding assays.
543-544	14	Comment: Does this requirement also apply to chromatograms from animal PK studies?	The text has been revised and indicates now that for

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		5 to 20 % of chromatograms of subjects is far too much to include into a study report (e.g. for a population PK study 20% of subjects could mean: 200 subjects with 5 blood sampling time points resulting in 1000 chromatograms for the subjects + 60 QC chromatograms + 200 calibration standard chromatograms). Some examples of chromatograms should be sufficient and additional chromatograms could be delivered in an electronic format on request if needed. **Proposed change:** To ask only for chromatograms of subjects from bioequivalence studies and not for chromatograms from other human and animal trials: "*For bioequivalence studies at least 5 % of chromatograms of the subjects including the corresponding QC samples and calibration standards should be included in the study report.*"	bioequivalence studies 20% of the chromatograms should be included and for other studies the relevant chromatograms can be appended and should be available on request.
543-544	15	Please distinguish between animal and human studies. A few examples as an integrated part of the report should be sufficient. 5 to 20 % of chromatograms of subjects could be delivered in an electronic format added to the electronic version of the bioanalytical study report. **Proposed change:** The bioanalytical report should contain some typical chromatograms as examples. Additional chromatograms can be attached to the electronic version of the report.	The text has been revised and indicates now that for bioequivalence studies 20% of the chromatograms should be included and for other studies the relevant chromatograms can be appended and should be available on request.
543-544	21	Clarification is needed as to whether this requirement applies for non-clinical or just clinical studies. **Proposed change (if any):** We recommend that only aa few (≤6) representative chromatograms would need to be added to nonclinical reports.	The text has been revised and indicates now that for bioequivalence studies 20% of the chromatograms should be included and for other studies the relevant chromatograms can be appended and should be available on request.
543-544	39	What is the reason for the requirement to include the	The text has been revised and indicates now that for

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		chromatograms from at least 20% of the subjects? This is simply the waste of paper, time and money with no benefits. A set of chromatograms from a single subject and a calibration curve is sufficient for everyone to judge the quality of separation, peak shapes and peak integration. The additional chromatograms represent only useless tons of papers, fuel burned for transportation of these papers, etc. **Proposed change (if any):** Reduce the required number of chromatograms significantly.	bioequivalence studies 20% of the chromatograms should be included and for other studies the relevant chromatograms can be appended and should be available on request.
543-544	45	Follow FDA recommendation for chromatograms	The text has been revised and indicates now that for bioequivalence studies 20% of the chromatograms should be included and for other studies the relevant chromatograms can be appended and should be available on request.
544	23	Including 5 to 20% of the subjects represent a lot of chromatograms. We suggest that it be necessary to include only representative chromatograms e.g. 1 subject/animal by sex and by group. Is this really of benefit in report interpretation, particularly when this may represent only a small part (phase report) of the overall study report?	The text has been revised and indicates now that for bioequivalence studies 20% of the chromatograms should be included and for other studies the relevant chromatograms can be appended and should be available on request.
545	24	Comment: The definition list is incomplete since there are important concepts, such as "analytical run" or “matrix factor” which definitions are not included, apart from many other definitions such as analyte, biological matrix, etc. This section should be complemented and harmonized with the Glossary section of the FDA Guideline.	The definition list has been updated.
545 Definitions	36	Comments It is proposed that the following definitions are added to the definition glossary Analyte: A specific chemical moiety being measured, which can be intact drug, biomolecule or its derivative,	The definition list has been updated.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		metabolite, internal standard, and/or degradation product in a biologic matrix. Analytical (or batch) run: A complete set of analytical and study samples with appropriate number of standards and QCs for their validation. Several runs (or batches) may be completed in one day, or one run (or batch) may take several days to complete. Internal standard: Test compound(s) (e.g. a structurally similar analogue, or stable isotope labelled compound) added to both calibration standards and samples at a known and constant concentration to correct for experimental variability during sample preparation. Nominal concentration: Theoretical or expected concentration. Recovery: The extraction efficiency of an analytical process, reported as a percentage of the known amount of an analyte carried through the sample extraction and processing steps of the method. Reproducibility: The precision between two laboratories. It also represents precision of the method under the same operating conditions over a short period of time. Sample: A generic term encompassing controls, blanks, unknowns, and processed samples, as described below: Quality Control Sample (QC): A spiked sample used to monitor the performance of a bioanalytical method in order to assess the integrity and validity of the results of the unknown samples analyzed in an individual batch. Unknown: A biological sample that is the subject of the analysis. Validation Sample (QC): A spiked sample, used to determine the performance characteristics of a bioanalytical method. In some instances a validation sample may be identical to a QC sample. Stability Sample A spiked sample, used to determine the	

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		stability of the analyte under the specified conditions. In some instances may be identical to a Validation (QC) sample. Stability: The chemical stability of an analyte in a given matrix under specific conditions for given time intervals. System suitability: Determination of instrument performance (e.g., sensitivity and chromatographic retention) by analysis of a reference standard prior to running the analytical batch.	
545-589	21	We would recommend to add clear definition for: 1. Full validation 2. Partial validation 3. Revalidation	The definition list has been updated.
545-589	45	Add clear definition for: 4. Full validation 5. Partial validation 6. Revalidation	The definition list has been updated.
545-591	26	Line 560: Carryover can occur with any concentration. It can be a system carryover, which is approximately independent of sample concentration. It also can be sample-dependent, where it is most pronounced at higher concentrations. Line 568: See comment for line 151. Lines 584-585: The definition of "Selectivity" is too vague to be useful. The phrase "which may be expected to be present" covers all possible sources of interference, including all listed in the definition of "Specificity". Typically "Selectivity" has been used to refer to the ability to distinguish the analyte from non-specific interferences from matrix components. Line 587-588: The definition of "Specificity" should not be restricted to ligand-binding assays, but should apply to chromatography-based assays also. This definition is also too broad. Typically, "Specificity" has been used to	The definition list has been updated.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		refer to the ability of the analytical method to measure only the analyte in the presence of biotransformation products, concomitant medications and other specific chemical components in the sample.	
550-551	37	Comment: Mathematical formula should be added: $\text{ACCURACY} = \frac{\text{determined value}}{\text{reference value}} \cdot 100\%$ **Proposed change (if any):** The accuracy of an analytical procedure expresses the closeness of the determined value to the value which is accepted either as a conventional true value or an accepted reference value (e.g. nominal concentration).	The definition list has been updated.
556	11	Comment: consider using "highest" and "lowest" for the "upper" and "lower", respectively.	The definition list has been updated.
558	11	Comment: add "the" after "meet"	The text has been revised accordingly.
561	11	Comment: remove "peaks"	The text has been revised accordingly.
565	24	Comment: In the Definitions section, the definition of incurred sample could be expanded to address the "incurred" aspect (e.g. retesting or reanalysis of actual study samples during subsequent runs).	The definition list has been updated.
565-567	37	Comment: this definition is unnecessary and may be deleted. **Proposed change (if any):** ~~Incurred sample reanalysis~~ ~~The analysis of a portion of the incurred samples to determine whether the original analytical results are reproducible.~~	This is not agreed.
569	23	Should this not be lower limit of quantification as in the previous line.	Agreed.
569-570	37	Comment: constant terminology should be used, additional sentence is proposed.	The definition list has been updated. The text has been revised accordingly.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): The lower limit of quantitation ~~limit~~ of an individual analytical procedure is the lowest amount of analyte in a sample which can be quantitatively determined with pre-defined precision and accuracy. The LLOQ is set by the lowest calibration standard.	
571-573	43	So called "Matrix effect" cannot be generally defined, predicted or investigated in detail as it is a consequence of many unidentified chemical reactions taking part in the ion source of the mass spectrometer. Generally it changes not only by the composition of each individual biological sample but also it changes by time and by the history of the measurements on the individual instrument (traces of the adsorbed strong Bronsted acids or bases in the ion source usually play much more important role than weaker acids or bases present and extracted from the biological matrix itself). In some cases of analytes and internal standards it is a real critical factor, in others it may be neglected and need not to be investigated at all. So quite an extensive modification of this paragraph is suggested and recommended including the definition itself. Next improved definition of the Matrix effect (rows 571-573) is suggested: Matrix effect is defined as either enhanced or suppressed ion formation from the analyte/internal standard caused by the competitive/parallel acido-basic reactions in the ion source. Undefined components extracted from the biological matrix and/or adsorbed in the ion source participate in this competitive set of reactions.	The definition list has been updated.
572-573	37	Comment: the definition is a bit misleading, because identified substance may also cause matrix effect **Proposed change (if any):** The direct or indirect alteration or interference in response due to the	The definition list has been updated.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		presence of identified or unintended interfering substances in the sample.	
574	24	Comment: Definitions should be presented.	The definition list has been updated.
581-582	37	Comment: This definition suggests that not linear response function is not acceptable, which is not scientifically justified. **Proposed change (if any):** Response function is a function which adequately describes relationship between instrument response (e.g. peak area or height) to the concentration (amount) of analyte in the sample. Response function is defined within a given range.	The definition list has been updated. The text has been revised accordingly.
582	11	Comment: in some cases, it may not be "directly proportional".	Agreed. The definition list has been updated.
582	45	"results which are directly proportional to the concentration (amount) of analyte" – it may be inversely proportional as well (e.g. in competition ELISA).	Agreed. The text has been revised.
584-585	37	Comment:- **Proposed change (if any):** Selectivity is the ability of the bioanalytical method to measure and differentiate the analyte in the presence of components which may be expected to be present in the sample.	The definition list has been updated.
585	11	Comment: add "other" before "components" and add "in a sample" at the end of this sentence	The definition list has been updated.
589	24	Comment:. **Proposed change (if any):** "upper quantitation limit" of to "Upper Limit Of Quantification" Additional comments Please add definitions for full validation, partial validation and re-validation	Agreed. The text has been revised. The definition list has been updated.
590-591	37	Comment: constant terminology should be used, additional sentence is proposed.	The definition list has been updated.

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
		Proposed change (if any): The upper limit of quantitation ~~limit~~ of an individual analytical procedure is the highest amount of analyte in a sample which can be quantitatively determined with pre-defined precision and accuracy. The ULOQ is set by the highest calibration standard.	
591	21	We would recommend to include a reference section with URL addresses where appropriate for the reference documents quoted as this would in particular help the understanding of Section 3 of the guideline.	This is not agreed as URL addresses can change.
591-	23	There is no Reference Section **Proposed change (if any):** Add a Reference Section.	A reference section is considered not needed.