



EUROPEAN MEDICINES AGENCY  
SCIENCE MEDICINES HEALTH

18 February 2010  
EMA/574487/2009

## Overview of comments received on draft appendix 2 to the guideline on the evaluation of anticancer medicinal products in man (CPMP/EWP/205/95 Rev.3) on confirmatory studies in haematological malignancies

Interested party (Organisations or individuals) that commented on the draft Guideline as released for consultation

Stakeholder no.	Name of organisation or individual
1	EFPIA
2	Mundipharma Research Ltd
3	MSD
4	PSI
5	Roche
6	EORTC



**1. GENERAL COMMENTS – OVERVIEW:**

Stakeholder No. (see coverpage)	General Comment (if any)	Outcome (if applicable)
1	<p>While the guideline describes in detail the challenges, difficulties and ambiguities that clinicians encounter when developing compounds in haematological malignancies, we have several comments raised consistently, which relate to <b>acceptable endpoints</b> and design of registration studies. These comments are considered critical with high impact on the eventual operation of the guideline and their implementation in practice.</p> <p>The guideline does not sufficiently distinguish and provides only limited guidance on acceptable relevant endpoints specifically for the two types of treatments considered in this draft. Therefore the following points should be addressed and incorporated in the final guideline:</p> <p>(a) acceptable endpoints for treatments with aim to palliate (effects on QOL, transfusions, disease related symptoms, etc) and</p> <p>(b) acceptable endpoints for treatments with aim to change the natural history of the disease (cure, disease stabilization)</p> <p>With the exception of chronic phase-CML (Section 4.1), this document does not mention the acceptability of overall response rate (incl. complete response and prolonged duration of response, molecular response) as an endpoint in a registration study. This should be considered as a potentially relevant primary endpoint, especially in clinical settings:</p> <ul style="list-style-type: none"> <li>• Where little or no treatment options exist such as relapsed/refractory disease settings where patients have failed prior therapies</li> <li>• Where the treatment strategy is complex e.g. multiple myeloma where induction treatment is followed by transplant, is followed by consolidation and maintenance treatment – in these situations complete</li> </ul>	<p>Not understood. Please refer to proper sections of the appendix where these issues were discussed and are discussed in the revised document.</p>

Stakeholder No. (see coverpage)	General Comment (if any)	Outcome (if applicable)
	<p>response should be considered as primary endpoint for registration in early segments of the treatment, especially when CR/VGPR is correlated with long-term outcome,</p> <p>In addition, the terms of "PFS" and "DFS" are used interchangeably in the draft guidance, which leads to confusion. The definition of DFS and PFS are distinctly different in haematological settings and cannot be used interchangeably.</p> <p>In many haematological malignancies standard next line therapies might not be approved, but have been established through standard clinical practice. In these situations these commonly used therapies - although not registered - should be acceptable. Applicants may be advised to seek scientific advice on these aspects of the trial design.</p> <p>Use of active next-line evidence-based standard therapies, which are not necessarily licensed, could be acceptable and the impact of this on the possibility of detecting differences in OS as well as symptoms related to disease progression should be acknowledged</p>	<p>Please refer to the definition of "DFS" in the draft and the revised wording and re-naming ("EFS") in the revised document. DFS and PFS were not used interchangeably.</p> <p>Approval is not critical, but the evidence supporting the use of a certain regimen. If only established "in clinical practice" superiority in terms of efficacy has to be demonstrated for the experimental regimen.</p> <p>Acknowledged</p>
1	<p><b>Other major comment</b> We recommend inclusion of disease specific information on Chronic Lymphocytic Leukemia and Non-Hodgkin's lymphoma. (Proposal included at the end of the document).</p>	<p><b>Not acceptable</b> Major new sections cannot be included without providing stakeholders with the opportunity to comment. Will be considered at time of revision.</p>
4	<p>There are recommendations from an American Society for Haematology/US Food and Drug administration workshop on clinical endpoints in multiple myeloma which aims to provide "guidance, consensus and consistency in the definition of clinically relevant end points that can expedite new drug approvals". It would be useful if the EMEA could consider adopting these same definitions or at least acknowledge these existing guidelines for multiple myeloma trials in order to harmonise international trials. Reference: Anderson KC et al. Leukaemia (2008) 22, 231-239.</p>	<p>We have consistently avoided referring to specific consensus documents or guidelines. Instead sponsors are advised to consult up-dated and generally acknowledged documents.</p>

Stakeholder No. (see coverpage)	General Comment (if any)	Outcome (if applicable)
4	It would be nice to see the term “patient reported outcome” as well as “quality of life”.	It is acknowledged that “PRO” is gaining in popularity. “QoL”, however, is used in this appendix as in the main document.
5	While it is stated that PFS is acceptable, DFS appears to be the preferred endpoint throughout this document when cure is the aim of therapy. However, at least for diseases which have a solid tumor component (ie CLL, NHL), this is not considered appropriate and deviates significantly from the current standard, PFS. The inherent problem with DFS as an endpoint is that is affected by assessment of residual masses which could be classified as tumor or scar. Even when using PET or biopsies false positive and false negative cannot be ruled out. PFS is not affected by this and should remain the preferred endpoint in these conditions. EMEA may consider that if DFS had been applied as a primary endpoint, not a single study used for filing purposes in the last 10 years in aggressive NHL would have been acceptable. Similarly, in non-curative intent studies, such as indolent NHL or CLL, PFS is questioned in the document as a valid endpoint, and DFS or OS is asked for. The outcome would be very similar, in that studies would need to be significantly increased in size and may no longer be feasible due to the low incidence of these diseases and multiple ongoing competing projects. In some diseases such as MDS the use of a OS end-point would also not be appropriate if not for evaluating detrimental effect because of the age of patients that are usually in their 70ies and therefore close to the median survival of mankind.	Agree, PFS may be used in high grade lymphoma. With respect to studies conducted with non-curative intent, DFS was <u>not</u> mentioned. Instead and as proposed PFS is the preferred measure.  With respect to MDS, PFS is mentioned as being an end point of relevance while survival is viewed as a safety endpoint in low risk MDS.
5	Genetics and molecular markers use are also not currently included as part of the guideline – indeed the use of molecular response is argued against for CML – which will make the guideline itself outdated very soon. With regard to CML, it could already be considered outdated in this respect.	With respect to CML and “molecular response” it may be debated whether it is sufficiently validated yet to be used as primary end point. In general, however, use of molecular markers, e.g. defining minimal residual disease, is encouraged, but not as primary end point until further validated
6	Abbreviations should be defined when firstly indicated (e.g. BSC, OS, PFS, NfG, etc)	Accepted
6	Terms like DFS, EFS, PFS are extensively used in this document with no clear definition. It must be acknowledged that recommendation of a clear definition of	

Stakeholder No. (see coverpage)	General Comment (if any)	Outcome (if applicable)
	<p>these endpoints would avoid variability in reporting trials results.</p> <p>In this respect, the issue on secondary neoplasms should be included as well. For the EFS/DFS/PFS endpoints those who died without relapse are considered as events at time of death. Among them, there are patients who developed a second neoplasm. Some non-EORTC Groups consider them as events at the time of occurrence of the 2<sup>nd</sup> neoplasm. What is the EMEA position in this respect?</p> <p>In case several competing risks may influence the outcome (e.g. progression and death without progression), even if the primary endpoint is PFS for the treatment comparison, the respective sub-endpoints should be analysed (e.g. cumulative incidences of progression and of death without progression).</p>	<p>For the primary analysis of PFS, occurrence of a second neoplasm should not be regarded as an event if not specifically justified, but could be regarded as an event in an event-free survival analysis.</p> <p>Agree, of relevance, e.g. if death without progression constitute a relevant proportion of events.</p>

## 2. SPECIFIC COMMENTS ON TEXT

Line No.	Stakeholder No.	Comment and Rationale; proposed changes	Outcome
1-2	1, 5	<p><i>“The aim of this appendix is to provide guidance on the design of confirmatory studies in patients with haematological malignancies.”</i></p> <p>Please clarify definition of confirmatory studies (i.e., ICH E9 versus CPMP/EWP/205/95/Rev. 3/Corr.”) and specify whether the definition of confirmatory studies is different for the different types of MAA approvals, i.e., Conditional, Exceptional, and Normal.</p> <p><b>Proposed change (if any):</b></p>	Lines 6 – 11 provide a sufficiently precise definition.
9-11	1, 5	<p><b>Comments:</b> Proposal for change/addition:</p> <ul style="list-style-type: none"> <li>- Alignment made with line 51-52 of draft guidance which states BSC is formally accepted.</li> <li>- Although confirmatory studies are generally required to be randomised, reference-controlled in nature, it should be acknowledged that this may not always be feasible depending on the disease setting and the treatment intent. This is in line with the main guidance document section <i>III.2.7 Use of external control</i>: which acknowledges that prospective confirmation in randomised, reference-controlled studies is not only unacceptable to investigators, patients and ethics committees, but also unnecessary in situations where the treatment effect is dramatic and the usual course of the disease highly predictable.</li> </ul> <p><b>Proposed change:</b> <u>In general</u>, these studies are randomised, reference-controlled in nature and the target population, as well as the reference regimen (<del>may be</del> including BSC), are normally defined by disease, stage and prior lines of therapy.</p>	<b>Partly accepted:</b> “In general” included.

		<p><u>There may be however disease specific exceptions.. Further in those occasional situations where the treatment effect is dramatic and the usual course of the disease highly predictable, it is also acknowledged that prospective confirmation in randomised, reference-controlled studies is not only unacceptable to investigators, patients and ethics committees, but also unnecessary in situations where the treatment effect is dramatic and the usual course of the disease highly predictable.</u></p>	<p>As correctly stated, paragraph III.2.7 of the main document covers the use of external control. This is considered sufficient and a reference to the main guideline has been included in the “Scope”.</p>
12 - 15	1	<p><b>Proposed addition:</b> Proposal for addition:</p> <p>In many haematological malignancies standard next line therapies might not be approved, but have been established through clinical trials.</p> <p>For instance, use of active next-line <del>evidence-based standard therapies, which are not necessarily licensed</del>, must be accepted and the impact of this on the possibility of detecting differences in OS and other time related efficacy parameters as well as symptoms related to disease progression should be acknowledged.</p>	<p><b>Not accepted:</b> The main document (III.1.4) provides a definition of “evidence- based” and clarifies that licensure is not a prerequisite.</p>
20 - 24	1, 5	<p>Proposal for clarification:</p> <p>The sentence starting "For haematological malignancies where treatment is administered without curative intent" is confusing and this should be rewritten in a simpler more summarized form.</p> <p><b>Proposed change:</b> For haematological malignancies where treatment is administered <del>without curative</del> <u>with palliative intent</u> the choice of comparator and primary endpoint should be based <u>on the expected efficacy and toxicity profile of experimental treatment. Sponsors are advised to seek scientific advice on such aspects of the trial design.</u> <del>there are often alternative, in clinical practise well established regimens, showing major differences in efficacy and toxicity indicating that efficacy in these cases is</del></p>	<p><b>Not accepted:</b> This appendix differentiates between treatments administered with “curative intent” and “non-curative intent”, further subdividing the latter: if the intent is to achieve long-term disease control or palliation. This paragraph refers to “without curative intent”.</p>

		<p><del>considered to parallel toxicity. It is therefore of relevance in the planning phase, to take into account the expected tolerability/toxicity profile of the experimental regimen compared with the selected reference regimen.</del></p>	
29 – 30	1	<p>Proposal for clarification:</p> <p>The sentence starting with “If, however, exploratory study data ...” is not clear and seems to imply that Overall Survival would still be strongly recommended as the primary endpoint even in situations where PFS would be sufficient for licensure. This is rather ambiguous and contradictory to the statement in line 15-16. A statement in-line with other parts of the guidance, with allowance for PFS but accompanied by a significant body of data on OS would be more logical.</p> <p>In addition, it is understandable that superiority trials might be preferred over non-inferiority trials by regulatory authorities, however a strong recommendation seems to discourage development of compounds that could provide similar or reduced toxicity and/or lead to substantially improved quality of life with similar efficacy compared to the standard of care.</p> <p><b>Proposed change:</b> If, however, exploratory study data indicate that, e.g. a <u>superior activity in terms of PFs and OS survival benefit</u> is a realistic aim also for a compound with similar or reduced toxicity, <del>it is strongly recommended that the a study is designed to demonstrate superiority should be considered this,</del> even in situations where non- inferiority <del>in terms of PFS</del> would be sufficient for licensure.</p>	<p><b>Not accepted:</b> If exploratory data indicate that a survival benefit is a realistic aim, it is in the best interest of patients, the health care system and the companies that the confirmatory study is designed to show this. This does not imply that survival should be the primary end point but the integrity of the trial should not be jeopardised by too early, non-inferiority PFS analyses.</p> <p>This is <u>not</u> discouraging the development of compounds with similar efficacy and improved safety and is not in contrast to what is said in lines 15-16.</p>
29 – 32	5	<p>The sentence starting with “If, however, exploratory study data ...” is not clear. If an experimental drug does not improve PFS it is highly unlikely to improve survival. The only exception could be a supportive care drug that does reduce, says, infections but in</p>	<p>See above. A high CR rate or durable responses in exploratory (single arm) studies could indicate that survival can be favourable affected. Even if the experimental compound show good tolerability it would be unwise to aim for minimal passing level, e.g. non-inferiority in terms of PFS something that would</p>

		this case it should be clearly spelled out.	be possible to show with only few patients (and a favourable point estimate).
35 - 39	1, 5	<p>The reference to SAEs is misleading and inappropriate because this category of events includes events such as pregnancies that are not per se toxic events.</p> <p><b>Proposed change:</b> Major increase in toxicity”, however, in most cases refers to a fear that treatment-related <u>irreversible adverse events or severe impairment to patient condition</u> SAEs will be relevantly increased in the experimental arm.</p>	<b>Accepted.</b>
33	5	<p>In a typical registration trial of A is compared to A + B (eg FC vs R-FC in the CLL studies). Maintenance vs observation trials are other examples. It is very likely that more SAEs occur in the experimental arms in such designs, but this is not synonymous with a ‘major increase in toxicity’.</p> <p>Reconsider safety categories. Differences on a 5% basis or so might be more relevant than any absolute difference or at least explain need for an appropriate benefit/risk assessment on case by case basis.</p>	<b>See above</b>
37 – 38	1	<p>Proposal for deletion:</p> <p>The need for prophylaxis should not be regarded automatically as a major increase in toxicity. Such events that are adequately managed with appropriate prophylaxis should be considered as such and therefore should not be regarded as major toxicities. Instead an evaluation of the overall risk benefit of the treatment should be done.</p> <p><b>Proposed change:</b> <del>However, in a comparison with a well tolerated reference regimen of low toxicity, for example, sustained immune suppression with infectious complications and need for prophylaxis could also be regarded as a major increase in toxicity.</del></p>	<b>Accepted</b>

43 – 46	1, 5	<p>Proposal for rewording, additional guidance and moving of paragraph:</p> <p>The term “children” should be exchanged with "for the paediatric population (age 0 to 18).</p> <p>The timing of long-term collection of toxicity should be considered in light of the median life expectancy. In some cases, there might be very rare examples of long-term survivors such as in relapsed ALL. It is impractical as well as of limited scientific value to collect data from single or very rare cases in such instances. Guidance as to how best collect long-term safety should be provided such as secondary malignancies, for which the information is generally not possible at time of registration. In general, a 5-year follow-up period is considered sufficient in adults for collection of long term safety data. In children, with the exclusion of secondary tumors later in life, most of the long-term events also become apparent in this time period.</p> <p>In addition the paragraph should be moved before the paragraph starting at line 77.</p> <p><b>Proposed change:</b> For <del>children</del> <u>the paediatric population (0-18 years)</u>, toxicity data should be considered with special emphasis on long-term toxicity and in relation to projected <u>median</u> life expectancy and risks of interference with QoL. Usually a 5-year period is considered sufficient in adults. In children, with the exclusion of secondary tumors later in life, most of the long-term events also became apparent in such time period.</p> <p><u>Long-term safety in children could be handled as post approval commitment in situations where a randomized trial showing a superiority of the experimental product forms the basis for registration.</u></p>	<p><b>Accepted</b></p> <p><b>Not acceptable</b> as such. Long-term survivorship is the rule rather than the exception in haematological malignancies in the paediatric population (including ALL, AML, HL and NHL; e.g., Pui 2006, Kaspers 2007, Pulte 2009, Reiter 2007). Only part of relevant adverse events present within 5 years after diagnosis. Some relevant adverse effects may only become evident at developmental milestones or when adversely affecting growth, psychomotor, psychosocial or sexual development, for example. Other adverse effects may become apparent in special situations, such as heart failure due cardiotoxic medicines when giving birth. For some of haematological malignancies in the paediatric population, outcome with respect to long-term treatment-related toxicities is becoming as important as maintaining the outcome with respect to the disease.</p> <p><b>Partly accepted.</b> The paragraph will be rephrased to reflect the general need for data on long-term toxicity in the paediatric population at the latest when confirmatory studies are conducted, however leaving it open how to do the follow-up.</p> <p><i>Please note that paediatric sections have been removed from the document and will be included in the upcoming revision of the Addendum on Paediatric Oncology (EMEA/CPMP/EWP/569/02).</i></p>
43 – 46	5	<p>While it is acknowledged that QoL is important, measurement of QoL and inclusion in labels has not been straightforward over the years – one reason for this is use and acceptance of appropriate</p>	<p>Several tools for assessing health-related quality of life are already available, and if not available, may need to be adapted or developed. For example, the PEDQoL is validated for an</p>

		<p>tools. Further guidance on this could therefore be considered,</p> <p>Further guidance on appropriate tools and measurement of QoL as well as clearly separating this out from safety/AEs</p>	<p>increasing number of countries / languages, and a number of other instruments are available.</p> <p>Not accepted, as this is beyond the scope of the guideline. QoL may well be used for benefit – risk evaluation, in addition to efficacy and toxicity data.</p>
45 – 46	2	<p><b>Comments:</b> “Toxicity in respect to impediments of organ- and neurological and <b>psychosocial</b> development should be under special focus.”</p> <p>We have concerns that the term “psychosocial development” is unclear and open for further discussion. We would wish to learn more about the expectations of the EMEA and where relevant the PDCO.</p> <p>Causality is unclear due to the subjective nature of the condition.</p> <p>This has a potential impact on a number of factors: the type of study required to demonstrate psychosocial development, e.g. large scale epidemiology; duration of the study (possibility of &gt;20 years); the insurance of the trial and indemnity periods; reporting intervals; how to demonstrate commitment to follow-up, particularly with long term survival; risk/benefit impact of survival versus psychosocial development; impact on PIP, particularly with regard to deferrals, PIP compliance and subsequent rewards. We would request confirmation of who is responsible for addressing this question – the EMEA or PDCO.</p>	<p>The term “psychosocial development” is defined in paediatric medical science and published literature.</p> <p>Not understood.</p> <p>The concerned paragraph will be rephrased to reflect the general need for data on long-term toxicity in the paediatric population by the time confirmatory studies are conducted, however leaving it open how to carry out the follow-up, as this also depends on the medicinal product or its mechanism of action. Proposals for the long-term follow-up are part of applications for agreement of paediatric investigation plan applications.</p> <p>Discussion of the Paediatric Regulation is beyond the scope of the guideline.</p> <p><i>Please note that paediatric sections have been removed from the document and will be included in the upcoming revision of the Addendum on Paediatric Oncology (EMEA/CPMP/EWP/569/02).</i></p>

47 – 54	1	<p>Where there is no single standard of care, investigators usually choose the best treatment among several based on different considerations (patient characteristics, objective of treatment). These chosen treatments may differ in toxicities. The current text in this paragraph states that the chosen control regimen should be one that is considered a first choice in clinical practice, and one with similar expected toxicity to the experimental therapy. This however, may not be always possible and as outlined will depend on investigator preference and the main objective of the study, e.g. showing superior efficacy or better tolerability and/or QoL.</p> <p>Proposal for reference to the main guidance should be included, since the main guidance contains additional important information with regard to “investigators choice” including required study design and statistical approach.</p> <p><b>Proposed change:</b> <u>Where multiple options exist, consideration should be given to the toxicity and efficacy profiles of the available reference regimens, to ensure that the selected reference regimen will provide an appropriate comparator for the assessment of the benefit-risk of the experimental agent Further information is available in the main guideline. (“Guideline on the evaluation of the anticancer medicinal products in man”, CPMP/EWP/205/95/Rev.3“, section III.1.4)</u></p>	<p><b>Partly accepted:</b> A general sentence has been included in “the scope” stating that the main guideline text should be consulted, e.g. with respect to reference regimen/investigators best choice.</p> <p>Current wording: “The benefit – risk of the reference regimen should be well documented and the regimen should be considered a first choice in clinical practice. Among such regimens, a regimen with similar expected toxicity to the experimental regimen is preferred if available and <u>suitable from a design perspective.</u>”</p> <p>“Suitable from a design perspective” was meant to cover situations where the objective was to improve tolerability without loss in efficacy. Therefore “including the objectives of the trial” has been added.</p>
47 - 54	5	<p>Restricting the choice of the comparator regimen to the ones with similar toxicity will defy the purpose of increasing tolerability. If a regimen has for example Neutropenia as main toxicity it would be more appropriate to evaluate it against a regimen that does not induce Neutropenia to evaluate the risk/benefit of affecting the disease as well different target organs rather than stick to the same toxicity pattern.</p>	<p><b>Accepted.</b> See above “including the objectives of the trial”.</p>
52	5	<p>Further clarification re: the use of investigator’s choice as a comparator when evaluating toxicities. Using investigator</p>	<p><b>Not accepted,</b> but wording slightly revised: BSC is acceptable in these cases, but an active comparator, documented e.g. in</p>

		choice may lead to many different toxicities with different regimens. This may be particularly relevant in cases where “cure” is not a goal of therapy. One could foresee a regimen which might be less toxic than one comparator in renal side effects but more toxic in cardiac side effects, and the numbers could in the end be too small for any definitive conclusions. This would be particularly relevant in the “non-inferiority” setting.	terms of response rate, is often preferable. If a single reference regimen cannot be defined, investigator’s best choice is an option.
52	5	While investigator choice may be feasible in some settings, in many cases it is still not feasible (refractory settings) or has its own inherent issues (such as variability across regions of different drugs). Therefore, this should also be acknowledged and explained.  Clarification that BSC is still acceptable in some settings/investigators choice has its own issues to consider	It is acknowledged that there are problems also with the “investigator’s choice” option. However, variability across regions should be possible to handle as this approach provides some freedom.  <b>Partly accepted:</b> Current wording “Formally, BSC is acceptable in these cases”. New wording: “BSC is acceptable in these cases, ..”
54	1, 3, 5	The advice to not cross-over patients to the experimental treatment group may not be in the patients' best interest when a new and more efficient or more tolerable medicine is discovered. Especially in the refractory setting cross-over after progression is usually a pre-requisite for patients and investigators to participate in such a study. It should be acknowledged similar to the main guidance that cross over to experimental arms may happen in these situations. Guidance should be provided re statistical aspects in cases where this cannot be avoided.  <b>Proposed change:</b> <u>While cross-over to the experimental arm should be avoided, it is acknowledged that in certain situations this may be in the patients' best interest.</u>	<b>Not accepted.</b>  In the main guideline it is stated: “It is thus recognised that investigators, patients and ethics committees may require, e.g. optional cross-over at time of tumour progression.”  In the appendix it is stated. “Cross-over to the experimental arm should be avoided”.  In the last-line setting, the appendix prioritises the use of investigator’s best choice over BSC. The reasons for this are twofold. It enables a comparison with currently used treatment options and reduces the pressure to cross-over patients to the experimental arm.  If the experimental therapy was reasonably well tolerated from an oncology perspective and if PFS was relevantly prolonged, e.g. 3 or more months compared with BSC, this would in most cases be seen as a positive outcome from a regulatory benefit – risk perspective and cross-over would constitute no major regulatory issue.

			<p>Such results, however, are uncommon in the last line setting and comprehensive data, including survival not confounded by cross-over, may be needed for a qualified benefit – risk assessment.</p> <p>Whether cross-over is in the patient’s best interest is prior to study results only an assumption, even though it is acknowledged that enrolment to the study is likely to be impaired if cross-over is no option, especially in a BSC study.</p> <p>There are no known proper ways to compensate for cross-over in the statistical analyses.</p> <p>Altogether it is found appropriate to retain the current wording: “Cross-over to the experimental arm should be avoided”.</p>
55 – 70	5	<p>Guidance is needed as to the use of molecular endpoints. Targeted therapeutics with reasonable toxicity profiles may increasingly use a molecular endpoint to suggest improved long-term survival/PFS/long-term symptom control. Given that molecular monitoring for CML is not yet validated, how should these less extensively studied endpoints be approached.</p> <p>Along this line of thought, the issue of Minimal Residual Disease as an endpoint to suggest improved long-term survival will likely become more prevalent in discussions re: CLL therapeutics as well as AML and CML. CR with MRD as a category will require validation of standards for MRD in CLL. Therapeutics in CLL should be aiming for MRD as long term studies will be need to address if this has an impact on long-term disease free survival. Given the recent data on the clinical significance of MRD on prolongation of PFS from Rituximab randomized FC vs FCR trial in first line CLL (Boettcher et al. ASH abs 326 and 3139) guidance is needed how the assessment and validation of MRD in CLL as endpoint should be approached.</p>	<p><b>Comment:</b> At this stage, use of molecular markers to estimate tumour burden is encouraged as secondary endpoints.</p> <p>In principle, validation of MRD as response variable poses no other problems than validation of ORR as surrogate for PFS or OS. Thus randomised comparative trials showing a difference in MRD response and PFS/OS are needed and the difference in PFS should be possible to explain based on the difference in MRD response. The prognosis of patients achieving MRD response should also be the same irrespective of treatment administered.</p>
55 – 62	1, 5	Treatment duration depends mainly on the mechanism of action, convenience of administration and toxicity profile of the product	

		<p>and the natural course of the underlying disease. Treatment duration should be justified taking all these factors into account.</p> <p><b>Proposed change:</b> <u>Treatment duration depends mainly on the mechanism of action, convenience of administration and toxicity profile of the product and the natural course of the underlying disease. Treatment duration should be justified taking these factors into account. For many solid tumours, treatment is administered until treatment failure, either due to tumour progression or unacceptable toxicity. This, however, is frequently not the case in haematological malignancies where a fixed number of cycles of therapy may be administered followed by watchful waiting, or, in some cases, allogeneic haematopoietic stem cell transplantation (HSCT). As the meaning of disease progression on and off therapy differs, this should be taken into account in the planning of studies with a fixed maximum number of treatment cycles and is of special relevance if the experimental regimen is more toxic than the control regimen and in case retreatment with the same regimen is an option after a sufficiently long period of time of non active disease</u></p>	<p><b>Not accepted:</b> If standard of care in a specific condition is a fixed number of cycles of therapy, treatment until disease progression should be demonstrated to be beneficial. In general terms, this cannot be inferred based on the sort of reasoning put forward by EFPIA.</p> <p>As the meaning of progression on and off therapy differs (as already stated in the appendix, e.g. with respect to resistance to further therapy), an endpoint beyond disease progression is necessary for an unbiased comparison.</p> <p>The current wording is unclear, however, and is therefore revised:</p> <p>“If standard of care encompasses a fixed number of treatment cycles, or treatment until stable response, for example, followed by watchful waiting and in order to support licensure for prolonged therapy, it should be demonstrated that this poses a favourable benefit risk. As the meaning of disease progression on and off therapy differs and as prolonged therapy might affect the activity of next-line therapies, these issues should be taken into account in the planning of such studies. If at all possible, these studies should be designed with the aim to document patient benefit in non-refutable terms, i.e. survival. Alternative endpoints, such as treatment-free survival, PFS on next-line therapy, etc. may be discussed in Scientific Advice procedures.”</p>
63 – 65	1	<p>Proposal for similar rewording to lines 55-62:</p> <p>Treatment duration depends mainly on the mechanism of action, convenience of administration and toxicity profile of the product and the natural course of the underlying disease. Treatment duration should be justified.</p>	<p><b>Not accepted,</b> see above (paragraph deleted covered by rewording of the former paragraph)</p>
66 - 69	1, 4	<p>From the last sentence, it is understood that "specific disease response criteria" can change with time and studies should be designed taking into account the most recent guidelines. The</p>	

		<p>guidance should be made clearer and more specific to state additionally that guidelines from the most recent international working groups at time of study start are acceptable.</p> <p><b>Proposed changes:</b> In contrast to solid tumours, it is acknowledged that disease specific response criteria are unavoidable in many cases and that full harmonization has yet not been accomplished for some disease entities. Therefore it is of importance to follow the progress made by international working groups on these issues. <u>It is acknowledged that usually versions of the guidelines may change over the course of the conduct of studies, in these situations the current version at time of study start is acceptable.</u></p>	<p><b>Accepted</b> (with minor modification): <u>It is acknowledged that usually versions of the guidelines may change over the course of the conduct of studies. In these situations the current version at time of study start is normally acceptable.</u></p>
66 - 69	5	Can examples or further clarification be given?	See above.
70 - 73	1, 5	<p>It is against the spirit of the orphan and new medicines legislation, to advise not to initiate studies in orphan indications until the benefit-risk is established in indications allowing for a more comprehensive evaluation. , especially with respect to safety.</p> <p><b>Proposed changes:</b> Some possible target indications comprise very small groups of patients, so small that “exceptional circumstance” might apply. <del>As a general recommendation, sponsors are advised to initiate studies in these patient groups only when benefit – risk is established in indications allowing for a more comprehensive evaluation, especially with respect to safety.</del> In these small target populations all evidence with respect to efficacy and safety must be taken into account. This encompasses outcome measures currently viewed as supportive only, such as HSCT rate, use of minimal residual disease to define response rate and recurrence of disease.</p> <p>Since 20 November 2005, ‘conditional approval’ was introduced</p>	<p><b>Not accepted:</b> The aim of the legislation is to make available reasonably well-documented medicinal products to patients with orphan diseases. For very small target populations, this means that circumstantial evidence derived from other indications may be the only way to gather essential information, i.e. to fulfil the aims of the legislation.</p>

		<p>to facilitate earlier access of innovative medicinal products to patients in the EU there exist possibilities to grant a market authorization “under exceptional circumstance” or to issue a conditional approval – these should be considered in these circumstances and sponsors should be advised to seek scientific advice.</p> <p><b>Proposed changes:</b> <u>In situations, where “conditional approval” is being considered by marketing authorization applicants, CHMP advice should be sought.</u></p>	<p><b>Not accepted:</b> Please refer to specific guidelines.</p>
74-76	1, 5	<p>Another area of concern is how to handle the many “orphan” indications within haematological malignancies. Many histological subtypes in haematology have incidences between very rare and rare and often represent severe conditions with few or no approved therapeutic options. In these low incidence indications, drugs showing clear signals of efficacy may have a therapeutic advantage . Because of inherent small patient numbers, fully powered comparative studies with time to event endpoints are not possible. Hence, evidence of efficacy can be provided by endpoints such as clinical response rates and duration of response, . Mature time to event endpoints such as PFS and even OS shall be important supporting secondary endpoints in these studies.</p> <p><b>Proposed changes:</b> <u>In cases, where histological subtypes of haematology malignancies occur at low incidences and which represent aggressive disease with few or no approved therapeutic options, drugs showing clear signals of efficacy may have a therapeutic advantage. Because of inherent small patient numbers, fully powered comparative studies with time to event endpoints are not possible. Hence, evidence of efficacy could be provided by endpoints such as clinical response rates and duration of response. Mature time to event endpoints such as PFS and even OS could be important supporting secondary</u></p>	<p><b>Partly accepted:</b> The comment is fully in line with what is already stated:</p> <p>“In these small target populations all evidence with respect to efficacy and safety must be taken into account. This encompasses outcome measures currently viewed as supportive only, such as HSCT rate, use of minimal residual disease to define response rate and recurrence of disease.”</p> <p><u>Reworded:</u> In these small target populations all evidence with respect to efficacy and safety must be taken into account. This encompasses <u>clinical response rates and duration of response</u> as well as <del>outcome measures currently viewed as supportive only, such as</del> HSCT rate, use of minimal residual disease to define response rate and recurrence of disease, <u>as appropriate. Mature time to event endpoints such as PFS and OS should be reported even though it is acknowledged that formal statistical</u></p>

		<u>endpoints in these studies.</u>	<u>significance cannot always be expected also if the experimental compound is relevantly active.</u>
70 – 76	5	<p>Agree that outcomes measures currently regarded as supportive such as HSCT rate, MRD, should be taken into account for "small target indications". However, it would also be good to allow this to be taken one step further for the future to allow for progress in developing these endpoints and as such also allow for endpoints such as MRD to actually be considered as primary in the future.</p> <p>Expand on wording to allow/support use of current supportive endpoints as primary endpoint in the future (assuming appropriate validation in place)</p>	There are many today exploratory endpoints (absolute tumour burden, change in absolute tumour burden, circulating tumour cells, tumour stem cell burden, etc.) which based on experience may serve as secondary endpoints and after proper validation as primary endpoint
70 - 76	5	<p>Some drugs in development will target only diseases with a specific mutation, and therefore will be useless in the general population by definition. Exposing patients to toxicities without any hope of benefit would be unacceptable.</p> <p>Suggest change the whole paragraph as such: “Some possible target indications comprise very small groups of patients, so small that exceptional circumstance” might apply. As a general recommendation, sponsors are advised to initiate studies in these patient groups PREFERENTIALLY when THE SAFETY PROFILE is established in indications allowing for a more comprehensive evaluation, especially with respect to safety UNLESS THE DRUG IS SPECIFICALLY DESIGNED FOR A DISEASE THAT AFFECTS SUCH SMALL NUMBER POPULATION.”</p>	<b>Partly accepted:</b> <u>Unless the target for activity is only expressed in these rare conditions and as a general recommendation, ...</u>
81	1	It is unclear why CML and AML would be preferred to NHL and ALL for extrapolation. Extrapolation between the adult and paediatric should depend on whether molecular and clinical course of the disease is very similar.	

		<p><b>Proposed changes:</b> In general, extrapolation from young adults may <del>more likely</del> be possible for CML and for AML, than for non-Hodgkin lymphomas or ALL. <del>However,</del> <u>and will depend on similarity of molecular markers, aetiology and, disease course and pharmacological action. seem similar</u> <del>However in these situations, in children and adults,</del> confirmatory studies <u>in children and adults,</u> are not <i>a priori</i> unnecessary, as there might be unknown, non-controllable factors.</p>	<p><b>Partly accepted.</b> CML and AML were essentially considered the (only) haematological malignant diseases where extrapolation may be possible, based on the similarity of biology, response and clinical course.</p> <p><i>Please note that paediatric sections have been removed from the document and will be included in the upcoming revision of the Addendum on Paediatric Oncology (EMEA/CPMP/EWP/569/02).</i></p>
82 – 83	5	<p>The following statement may need refinement: In general, extrapolation from young adults may more likely be possible for CML and for AML, than for non-Hodgkin lymphomas of ALL. Patte et al (Blood 2007) included in the largest ever study in pediatric NHL patients up to age 21, and the US COG group includes pts up to age 30 in pediatric protocols in NHL (Cairo M, ASH 2008)</p>	<p>Sentence will be rephrased to clarify. Paediatric oncology trials in the US regularly involve patients up to less than 22 years of age. The upper age range of trials with the paediatric population (defined as less than 18 years of age) in the EU may well be extended to 21 years of age, if scientifically justified.</p> <p>However, the mentioned studies were stratified by disease subtypes. As a result the various forms of NHL were differently distributed in the different age categories. Selection bias is more likely in older age ranges as compared to younger ages. Other age-related factors may influence outcome, for example, the diverging proportions of EBV and Burkitt lymphoma in different age groups.</p> <p><i>Please note that paediatric sections have been removed from the document and will be included in the upcoming revision of the Addendum on Paediatric Oncology (EMEA/CPMP/EWP/569/02).</i></p>
83-85	1	<p>Provided text is re-formulated to provide further clarity avoiding double negation.</p> <p><b>Proposed changes:</b> <del>However, even</del> When the aetiology, disease course and pharmacological action <del>seem</del> <u>are</u> similar in children and adults, confirmatory studies <u>may not be necessary</u> <del>are not a priori unnecessary. In this case justification should be provided as there might be unknown, non-controllable factors.</del></p>	<p><b>Partly accepted.</b> The concerned sentences were rephrased.</p>

84	6	<p>“confirmatory studies are not <i>a priori</i> unnecessary”</p> <p><b>Proposed changes:</b> confirmatory studies are required</p>	<b>Partly accepted.</b> The concerned sentences were rephrased.
87	1	<p>Added examples for disease specific guidance that is provided.</p> <p><b>Proposed changes:</b> For some individual disease entities (<u>i.e., CML, MDS, CLL, NHL</u>), more disease-specific guidance is provided</p>	Partly accepted: (i.e. CML, MDS, <del>CLL, NHL</del> ). This appendix does not provide specific guidance with respect to CLL, NHL. However, these entities may be covered by updates of the appendix if found reasonable.
90-92	1	<p>It is important to recognize that medical practice is one of the major considerations that need to be taken into account when designing confirmatory studies.</p> <p>Suggest referencing <u>current</u> "scientific evidence and/or generally acknowledged guidelines"</p> <p><b>Proposed changes:</b> With respect to diagnosis, criteria for initiation of treatment, eligibility, response criteria and choice of reference therapy, a justification based on <u>current</u> scientific evidence, <u>clinical practice</u> and <u>current</u> <del>generally acknowledged</del> disease guidelines is expected.</p>	<b>Partly accepted:</b> With respect to diagnosis, criteria for initiation of treatment, eligibility, response criteria and choice of reference therapy, a justification based on scientific evidence, and <u>updated and generally acknowledged</u> disease guidelines is expected.
96	4	<p>The guideline does a nice job covering various aspects of haematological malignancy trials; “stratification of the randomization for baseline covariates of major prognostic importance” in section 3 is particularly relevant. One of the important baseline covariates is total tumour burden, and fraction of ‘large’ lesions vs. ‘small’ lesions (since small tumours can grow fast and may be more susceptible to chemotherapy than larger ones), so it would make sense to stratify patients on this baseline variable.</p>	<b>Comment:</b> The appendix is intentionally left open with respect to covariates and only refers to “of major prognostic importance”. Tumour burden at baseline, for example, might well be of importance.
97 - 100	4	<p>Make it clear that adjusted analyses are always to be undertaken if stratification is employed</p>	<p><b>Agree</b> and in line with the points to consider document referred to.</p> <p><b>Revised text:</b> Whether stratification is undertaken or not, this should be discussed in the study protocol. In case adjusted analyses are to be undertaken for co-variates other than those</p>

			used for stratification, these factors should be pre-specified in the protocol or the statistical analysis plan (Points to Consider on Adjustment for Baseline Covariates CPMP/EWP/2863/99).
188 – 189 (sic)	1	The guidance should also give consideration to situations where at the start of a trial criteria may not be established, but they may evolve/develop over the course of the study – clearly these cannot be prespecified in the protocol.	<b>Comment:</b> This appendix covers confirmatory trials. Only exceptionally it is foreseen that essential “criteria” will be established over the course of the study and in those cases mainly based on data external to the study.
102	4	Paragraph starting line 102: Some of the sentences in this paragraph are not easy to understand	See comments to line 134 – 138 below.
102	5	Can examples be given of alternative criteria to define target populations than those commonly employed?	See comments to line 134 – 138 below.
109 – 113	1	Clarification of current text and adaptation to statistical practices. <b>Proposed changes:</b> <del>Whether stratification is undertaken or not, this should be discussed in the study protocol and in case adjusted analyses are to be undertaken, co-variables for such analyses should be pre-specified in the protocol. Primary analysis should be described in the protocol and justification of adjusting by stratification factor in primary or sensitivity analysis should be made. Any other covariates could be described either in the protocol or in the statistical analysis plan</del> (Points to Consider on Adjustment for Baseline Covariates CPMP/EWP/2863/99).	<b>Not accepted:</b> The message is that a <u>discussion</u> is warranted on these aspects in the protocol. See also above.
114 - 131	4	While PFS and OS are well recognized as the registration endpoints in haematological trials, there is a push in solid tumour therapy to use tumour load as a surrogate for OS. The Pharmacometric Division at FDA is advocating such an approach with an example in NSCLC, where they showed that baseline ECOG score, baseline tumour load and tumour load at Week 8 are reliable predictors of OS in phase III trials.	<b>Comment:</b> CHMP is in principle open for the use of surrogate/intermediate endpoints when properly validated. Use of and filing based on intermediate end points, however, might have consequences with respect to the integrity of the trial with respect to later end points such as PFS and OS. As surrogate end points rarely provide precise estimates of the effect on final endpoints, this must be taken into account and EU scientific advice is recommended.
116 - 119	1	Setting the time of these events at randomisation at time zero will not have the desired effect statistically. Censoring at time	<b>Comment:</b> This paragraph is of relevance essentially in the acute leukaemias and if treated with curative intent. The

		<p>zero excludes patients from the analysis. We suggest ordering events occurring during induction to reflect their importance - so set the time of failure as the time of death (for those occurring during as well as after induction) and for non-responders, set at either when the patient withdraws or at the end of induction if the course is completed as appropriate.</p> <p>The number of treatment cycles might not be mandated depending on the disease setting, therefore propose deletion of defined treatment cycles.</p> <p><b>Proposed changes:</b> There are conditions where efficacy failure is defined by not achieving <u>an appropriate response</u> (e.g., CR/PR) after a <del>defined number of</del> <u>certain treatment cycles (“induction”)</u> <u>and only patients with an appropriate response continue in the study.</u> In these cases non-<u>appropriate</u> response, relapse and death are counted as events in a disease-free survival (DFS) analysis <u>in an IIT analysis.</u> During the pre-specified induction time period, these events should be regarded as events at randomisation in the primary <u>IIT</u> analysis.</p>	<p>paragraph has thus been moved to 3.1. It was never the intention to censor patients at base line (“these events should be regarded as events at baseline”). However, the wording will be slightly modified. See below.</p> <p>Use of not achieving e.g. at least PR as an event in the primary analysis is discouraged. Rather these patients should be followed on or off next line therapy until disease progression or death.</p>
117 - 119	6	<p>Current wording: In these cases non-response, relapse and death are counted as events in a disease-free survival (DFS) analysis. During the pre-specified induction time period, these events should be regarded as events at randomisation in the primary analysis.</p> <p><b>Proposed changes:</b> In these cases lack of achievement of CR, relapse and death without relapse are counted as events in an event-free survival (EFS) analysis. Those patients who did not reach CR during the pre-specified induction time period, will be considered as events at time 0. Those who reached CR, will be considered from that time point at risk of relapse or death without relapse in the primary analysis</p>	<p><b>Accepted</b> and moved to 3.1, see comment above.</p>

116	4	Is the pre-induction period assumed to be pre-randomisation? If so, then the guidance makes sense but if this period is post-randomization then the guidance seems to go against intention to treat principles	See comments above.
118 - 119	5	Does this mean that the primary analysis should always be an ITT analysis no matter what? In case of rescue treatment after BMT or maintenance after induction the toxicity (early or late) of the BMT or of the induction regimen could overrule the efficacy of maintenance in responders despite a positive effect on survival. It depends from how many patients will die from the BMT itself, for example. It could be as high as 50% of them.	<b>Comment:</b> If the objective of the study is to investigate maintenance or rescue therapy, it is assumed that randomisation is undertaken after end of induction therapy/HSCT. Therefore the concern raised should be of minor importance. If the objective includes the consequences of HSCT/induction therapy, events related to HSCT/induction therapy should not be censored. Thus the ITT principle should be adhered to. If ITT for other reasons is considered suboptimal, regulatory Sc. Advice is recommended.
118 – 123	1	<p>- As outlined in major comments the terms DFS and PFS are used interchangeably in this guidance although they are distinctly differently defined in haematological malignancies.</p> <p>- The problem of how to determine "PFS" in leukaemia needs to be discussed in greater detail, particularly given that the determination of progression from anything but CR is very difficult to assess in acute leukaemia.</p> <p><b>Proposed changes:</b> In cases where PFS is as an acceptable primary endpoint for marketing authorisation, <u>PFS needs to be clearly defined and the consequences for the data collection on other endpoints must be carefully considered.</u> <u>and where the definition of PFS defers from the definition of time from randomization till recurrence or death for any cause, advice should be sought.</u> <del>and the</del> Consequences for the data collection on other endpoints must be carefully considered.</p>	<b>Comment:</b> This is not correct. Please notice that in the revised appendix event-free survival is used instead of disease-free survival to somewhat better adhere to the terminology used by the IWG.
124 – 126	1, 3, 5	When PFS is an acceptable endpoint, and has clinical significance, approval including a statement in the label that OS is not known should be granted. OS data can then be provided	

		<p>after marketing authorization.</p> <p>While it is acknowledged that overall survival data will be needed to assess the benefit-risk of the given treatment at the time of the PFS analysis, overall survival will be certainly affected by the cross-over to other treatments. In these situations, it would be preferable to use survival rates at given time-points to exclude a definite – negative – effect on survival.</p> <p><b>Proposed changes:</b> <del>Even When PFS is an acceptable the</del> <u>primary endpoint, and the effect shown is of clinical significance, approval is possible without overall survival (OS) data, however, the label will state that the effect on OS is not known. Sufficient data on overall survival OS or OS rates at specified time points can be provided subsequent to the original</u> <del>should sbe available at the time of marketing authorization.</del></p>	<p><b>Not accepted:</b> Survival data should be reported at time of submission, but it is accepted that these data may be immature. “Sufficient” was intentionally selected to give room for a justification taking the magnitude of PFS benefit, toxicity, expected survival etc. into account. OS is normally more informative than OS rates, but the latter may be used if specifically justified in the protocol.</p>
128	5	<p>While it is acknowledged that cross-over should be managed and minimized wherever possible it should also be acknowledged that this simply cannot be avoided in many situations. This should also be considered in conjunction with the clinical benefit and meaning seen with PFS as an endpoint in its own right.</p>	<p><b>Comment:</b> This appendix acknowledge the benefit of prolonged PFS in its own right.</p>
129	5	<p>In a non-inferiority study in children, PFS should still be sufficient. With OS sometimes in excess of 95% at 5 years, OS is not very meaningful, while reducing acute and long-term toxicity (cardiac) might well be. A PAC might be considered for long-term OS data.</p>	<p>Comment acknowledged.</p> <p><i>Please note that paediatric sections have been removed from the document and will be included in the upcoming revision of the Addendum on Paediatric Oncology (EMEA/CPMP/EWP/569/02).</i></p>
129	5	<p>EMEA document says: In children PFS is only suitable if an anatomical remnant without proliferative capacity is to be expected. This can be the case in, e.g. intrathoracic Hodgkin’s disease, intrathoracic B-cell lymphoma and bone localizations of lymphomas.</p> <p>PFS may be suitable in children with advanced refractory disease</p>	<p><b>Partly accepted.</b> Paragraph will be rephrased to clarify.</p> <p><b>Not accepted.</b> It should be defined whether the trial investigates</p>

		in which cure is considered unlikely and improvement in survival and/or palliation is still a worthwhile aim.	a medicinal product in a setting with the aim of a prolonged disease control or a palliation in the sense of the guideline. This distinction can be based on historical data. In general, for these distinct situations, the guideline applies similarly to the paediatric population as to the adult population.  <i>Please note that paediatric sections have been removed from the document and will be included in the upcoming revision of the Addendum on Paediatric Oncology (EMA/CPMP/EWP/569/02).</i>
129 – 131	1	Unclear why PFS is suitable if “anatomical remnant without proliferative capacity” is expected. Indeed this population can be considered cured. PFS should be generally allowed in advanced refractory disease settings. The age class should be specified if the word “children” is left.  <b>Proposed changes:</b> <del>In children PFS is only suitable if an anatomical remnant without proliferative capacity is to be expected. This can be the case in, e.g. intrathoracic Hodgkin’s disease, intrathoracic B-cell lymphoma and bone localizations of lymphomas. PFS may be suitable in children (0-18 years) with advanced refractory disease in which cure is considered unlikely and improvement in survival and/or palliation is still a worthwhile aim.</del>	Comment is similar to comment above.
132	1	<b>Proposed changes:</b> As frequently these studies cannot be conducted under proper double blind conditions, independent review of events of progression is <b><u>recommended</u></b>	See below.
132 - 133	5	IRC is expected - This will result in a considerable burden to patients due to multiple scheduled events and the necessity of ‘objective’ evaluations, such as CT-scans, as opposed to ultrasound or clinical exam. This inevitably leads to high radiation exposures that could at least be partly avoided. IRC in most haem diseases also not needed as diagnosis and response	<b>Partly accepted.</b> The burden to the patients of frequent imaging is acknowledged. It is also accepted that IRC review is meaningful only when progression mainly is based on imaging.  Even if OS is the primary end point, it is advised that the sponsor adheres to these principles. Frequently, there is only one

		<p>assessment largely based on objective lab data.</p> <p>Consider limiting IRC or at least also discuss the downsides of IRC in the document, ie increased radiation exposure, ‘objectiveness’ not always achieved, cost.</p> <p>Please clarify (eg. in line 132) that this paragraph (and therefore requirement for IRC) applies to confirmatory studies in which PFS is the primary endpoint. Roche also considers that an IRC is not required when PD is based on objective lab data or bone marrow evaluation, and that CT scans and Xrays should generally not be scheduled on a regular basis purely for detecting the first manifestation of PD in patients with haematological malignancies, unless this is part of routine clinical practice</p>	<p>pivotal study and PFS data are likely to be needed to support survival conclusions.</p> <p>Revised wording: As frequently these studies cannot be conducted under proper double blind conditions, independent review of events of progression <u>if based on imaging is recommended, also if OS is the primary end point. In large studies, it may be sufficient to undertake independent review in a justified proportion of cases.</u></p>
134 – 138	1	<p>Please provide guidance on design of confirmatory studies in molecularly defined diseases that may differ in their clinical characteristics and diagnosis (i.e. Jak2 mutant Polycythemia, Essential Thrombocythemia, Myelofibrosis). In order to adequately stratify for these different diagnoses in rare diseases, studies may become prohibitively large to conduct. Guidance is required for how to provide adequate analysis using targeted agents in these molecularly defined rare diseases.</p> <p>In addition, more clear distinction should be made between primary endpoint (e.g. overall survival time) and analysis method (i.e. 2 year event rate estimated according to Kaplan-Meier methods).</p> <p><b>Proposed changes:</b> Event rate at a pre-specified and justified point in time, e.g. at 2 years (<u>applying Kaplan-Meier methods</u>) might be used <u>as for the primary analysis of</u> the primary endpoint in these cases.</p>	<p><b>Comment:</b> This comment probably refers to lines 103 – 113, where it is stated that “... the study should be designed so that it is possible to conclude on the benefit – risk in the different subgroups of patients for which a claim is to be made.” At the same time this appendix acknowledges that for rare conditions, it is meaningless to require “prohibitively” large studies. As this must be addressed case by case, EU regulatory agreement is recommended prior to initiation of confirmatory studies (as already stated in the appendix).</p> <p>When survival rate at a single point in time is the measure selected for the primary analysis it is in most cases recommended that all patients should have been on study for that period of time and that the time point should be selected so that a clear majority of events in the long run should have occurred. (see below)</p>
136 – 137	6	<p>Current wording: “Event rate at a pre-specified and justified</p>	<p>Agree, see above. There are exceptions where the selection of</p>

		<p>fixed point in time, e.g. at 2 years might be used as primary endpoint in these cases.”</p> <p>Comment: What is important is that the expected event rate after a chosen time point is lower, far lower, than the one observed before that time point.</p>	<p>another point in time would be more sensitive to difference. This would require a justification in the protocol/SAP.</p> <p>For some conditions, events of progression will be observed at a slow rate making frequent assessments of events of progression a burden to the patients. Event rate at a pre-specified and justified fixed point in time might be used as the primary <del>endpoint</del> <u>outcome measure</u> in these cases. <u>When event rate at a single point in time is selected as the primary analysis, it is in most cases recommended that all patients should have been on study for that period of time and that the time point should be selected so that a majority of events in the long run should have occurred.</u> PFS should be reported as a supportive endpoint when a fixed time-point assessment is used as primary outcome measure.</p>
137 - 141	1	<p>Line 138 states that events falling between scheduled response assessments should be assigned to the end of the assessment interval in the primary analysis as an alternative. In addition, appendix 1 is referenced for further sensitivity analyses.</p> <p>Appendix 1 of the anticancer guideline states: "For the purpose of the primary analysis, the timing of a progression that is detected between two scheduled tumour assessments should be assigned based on the documented time of progression and not, for example, based on either scheduled time of assessment."</p> <p>Therefore, it is not clear which recommendation should be followed.</p>	<p><b>Accepted:</b> There is an inconsistency between appendix 1 and 2 on this point. The main message in appendix 1, however, is that alternative sensitivity analyses are warranted in this situation. Normally, assigning events to the end of the assessment interval would be conservative in trials aiming at showing a difference. The text, however, will be revised to be in line with appendix 1.</p>
138	6	<p><b>Comment:</b> This might be valid for studies where the event rate per time unit is low. Of course this should not be expanded to diseases where the event rate is high, where such an artificial delay of the event occurrence may introduce a bias in the treatment evaluation/comparison.</p>	<p>See above.</p>
143 -144	1	<p><b>Proposed changes:</b> The ultimate aim of therapy in patients with, e.g. acute leukaemia, <del>and</del> being suitable for intensive therapy is</p>	<p><b>Accepted.</b></p>

		to improve cure rate and survival.	
144 – 151	1	<p>The sequence of treatment in haematological malignancies can be extremely complex and could become even more in the future, e.g. induction treatment followed by high dose chemotherapy followed by transplant followed by consolidation followed by maintenance treatment. Due to this complexity there have been only few approvals in transplant settings in the EU, which is especially disappointing since transplant eligible patients are usually younger / fitter patients in need of better and especially curative treatments. The current wording of this paragraph is too narrow and does not provide ways of handling this complexity with regard to the design of confirmatory studies. This is especially valid for compounds studied for efficacy in early segments, such as induction before transplant, where response rate is a direct measure of activity, and in situations for which response rates are associated with long-term outcome.</p> <p><b>Proposed changes:</b> In some cases, however, and due to the complexity <u>and sequence</u> of administered therapies, the impact of an <u>relevantly active</u> experimental compound on <u>time-dependent</u> endpoints may be hard to demonstrate. For example, in case the experimental compound is used only as part of an induction regimen to be followed by consolidation therapy and possibly HSCT. It is foreseen that the experimental compound rarely will be used as single agent therapy, but will be used as add-on to an established, perhaps modified regimen, or as substitution for a compound being part of the established regimen. In this context, <u>treatment segments need to be defined and endpoints including response (e.g. CR) for each segment should be defined.</u> <del>maintenance therapy may be regarded as add-on therapy if maintenance therapy is considered non-established.</del> In case DFS/PFS <del>or response</del> is found to be a justified primary endpoint, it is of <del>special</del> importance that study data are analysed only when sufficiently mature. <del>, i.e. when it is foreseen that the DFS plateau is stable and cure rates can be estimated.</del> <u>Given the</u></p>	<p><b>Partly accepted.</b> This paragraph was meant as a preamble to set the scene and to highlight some of the complexities met in this setting, not to resolve all issues. However, there is agreement that due to the level of complexity and the many foreseeable scenarios that cannot be captured by this appendix, EU scientific advice should be considered in many cases.</p>

		<u>complexity of the situation, it is recommended that sponsors consider seeking scientific advice.</u>	
152	6	<b>Proposed changes:</b> DFS (time from CR till relapse or death without relapse) or EFS	Changed to EFS, see above. DFS as defined by IWG (and not as in the draft appendix) is informative, but in most cases ill suited as primary endpoint (cf. duration of response).
152		Reference is made to the major comment relating to endpoints and that PFS and DFS are used interchangeably in this guidance leading to confusion. Here the recommended endpoint should be DFS/EFS (for ALL and AML) to be in line with IWG criteria	See prior comments.
152 – 153	1	This paragraph suggests maturity is defined in terms of when the DFS curve plateaus. Given that DFS includes deaths unrelated to disease the DFS curve will never actually plateau and cannot be used to estimate cure rate; this will be particularly relevant in elderly populations. Suggest to define maturity in terms of the time after which additional <i>disease</i> recurrence is rare  <b>Proposed changes:</b> In case DFS is found to be a justified primary endpoint, it is of special importance that study data are analysed only when sufficiently mature, i.e. when it is foreseen that the DFS plateau is stable and cure rates can be estimated <u>or when additional <i>disease</i> recurrence is rare.</u>	<b>Accepted</b> , even though in the conditions of interest here, death due to other causes will be uncommon.
156	1	It would be desirable to specify that limited data collection on HSCT is acceptable in registration trials if the HSCT is part of the standard treatment administered to both treatment arms. The costs of the HSCT and the full data collection on HSCT can be prohibitive and may not add much to the experimental question.	<b>Comment:</b> This topic is of relevance when the experimental therapy is administered prior to HSCT and where due to toxicity an effect on EFS/OS should be demonstrated. As “difference” is the objective of therapy, it is in the sponsor’s interest to harmonise as far as possible indications for HSCT and procedures. However, full data collection is not considered necessary but the schedule for follow-up and event detection should be of high quality.
156 – 157	1	This section should provide clarity on how to deal with patients that receive transplant and if the agency accepts “bridging to HSCT” as acceptable surrogate endpoint.	<b>Comment:</b> The intention is to provide clarity. “Bridging” is not accepted a priori and should be further justified. As stated in the document: “As treatment administered prior to transplantation, for example, might affect outcome of HSCT, proportion of

			patients undergoing HSCT is not considered to be a suitable primary outcome measure...”
158 - 160	6	<b>Proposed changes:</b> It is fully acknowledged that criteria for HSCT (e.g. patient eligibility, HLA matching, conditioning regimen, GvHD prevention, etc) vary between institutions and regions. Nevertheless, these criteria and reasons for HSCT should be defined as well as possible in the protocol and should be captured by the CRF.	<b>Accepted</b>
161 – 167	1	<p>To include acceptability of CR as primary endpoint in transplant setting with secondary endpoints of PFS and OS in line with the changes and rationale proposed for changes in lines 144-151, i.e. The sequence of treatment in haematological malignancies can be extremely complex and could become even more in the future, e.g. induction treatment followed by transplant followed by consolidation followed by maintenance treatment.. The current wording of this paragraph is too narrow and does not provide ways of handling this complexity with regard to the design of confirmatory studies. Especially for compounds studied for efficacy in early segments, such as induction before transplant, where response rate is a direct measure of activity and in situations in which responses are associated with long-term outcome, patients are usually younger patients in need of better and especially curative treatments..</p> <p>It is agreed that censoring at HSCT is inappropriate especially when the rate of HSCT is high with regard to longer-term endpoints and using all patients in the analyses for DFS/PFS/OS is consistent with the ITT advocated in other oncology guidance as after all the fact that the patient responded well enough to receive HSCT was a benefit to the patient. Overall, it is favoured to apply no censoring regardless of the rate of HSCT, thus, “substantial” would also not need to be defined.</p> <p><b>Proposed changes:</b> <del>If the decision to transplant is purely defined by baseline characteristics, availability of donor, response or not</del></p>	<b>Not accepted:</b> The acceptability of a certain endpoint must be put in context. If there is a major increase in toxicity vs. similar toxicity. The paragraph about HSCT is focused on how to handle HSCT in the analysis plan, while appropriate endpoints are discussed in the following sections.

		<p><del>to therapy and</del> the proportion of patients transplanted is likely to be small, censoring at time of transplantation is acceptable in the primary analysis <u>as of response rate and DFS.</u> <del>, despite censoring, sufficiently well capture possible differences between study arms.</del> If, however, the proportion of patients transplanted is substantial, or if time to CR or quality of CR, for example, might influence the decision, censoring due to HSCT is considered inappropriate. <u>Where CR is validated for clinical benefit and can be reliably defined, CR can be used as primary endpoint similar to DFS or OS. should reported as primary endpoint without censoring for HSCT. When CR is used as primary endpoint DFS and OS should be reported as secondary endpoints. Given the complexity of the situation, it is recommended to have a case-by-case discussion with agencies.</u></p>	
164 – 167	3	<p><i>Re: If, however, the proportion of patients transplanted is substantial, or if time to CR or quality of CR, for example, might influence the decision, censoring due to HSCT is considered inappropriate. In these cases DFS or OS should reported as primary endpoint without censoring for HSCT.</i></p> <p>This specific guidance is confusing because HCST prospective patients should be excluded from the study. There should be a definition of how to handle patients who eventually go on to have a transplant.</p> <p>Should specify the labelling options or consequences for each of the potential endpoints (CR, PR) used in lieu of OS rather than mandating that DFS or OS be used without censoring for HSCT.</p>	<p><b>Comment:</b> The intention of the paragraph is to provide guidance with respect to handling of HSCT in the analysis plan. Further guidance with respect to proper endpoints is provided in the following paragraphs.</p>
161 - 164	6	<p><b>Proposed changes:</b> If the decision to transplant is purely defined by baseline characteristics, availability of donor, response or not to therapy and the proportion of patients transplanted is likely to be small, censoring at time of transplantation is acceptable for the primary analysis, <u>as, despite censoring, efficacy endpoints (response rate, EFS, DFS) may sufficiently well capture possible differences between study arms.</u></p>	<p><b>Accepted</b></p>

171	6	<b>Proposed changes:</b> as compared to adults, the donor ...	<b>Partly accepted.</b> Paragraph will be rephrased.  <i>Please note that paediatric sections have been removed from the document and will be included in the upcoming revision of the Addendum on Paediatric Oncology (EMA/CPMP/EWP/569/02).</i>
172 – 174	1	<b>Proposed changes:</b> Due to the lower complication rate in children and longer life expectancies as compared to adults, the donor sources of stem cells, <u>with</u> respect to HLA matching and donor-patient relation, are highly variable in childhood HSCT. Separate analyses <u>with respect to</u> the (projected) type of donor source should be done. This should also be considered <u>for</u> adult patients.”	<b>Partly accepted.</b> Paragraph will be rephrased.  <i>Please note that paediatric sections have been removed from the document and will be included in the upcoming revision of the Addendum on Paediatric Oncology (EMA/CPMP/EWP/569/02).</i>
175	1	Proposal for re-wording and re-structuring/re-naming:  Sub-title refers to “3.1 Treatment administered with curative intent” but not “haemopoietic stem cell transplantation”, which is currently the case.	<b>Not accepted.</b> Paragraph 3.1 is meant to provide guidance when treatment is administered with curative intent, whether HSCT is part of the treatment or not.
176 – 182	1	Statements in the 2 <sup>nd</sup> and 3 <sup>rd</sup> sentence seem to contradict each other. It is recognised that the contribution of a given treatment to a regimen is hard to define, presumably this comment is referring to situations where there is an absence of randomised trials, with DFS as an endpoint, comparing the comparator regimen with and without the substituted endpoint. However the 3 <sup>rd</sup> sentence then states ‘establishing absolute efficacy’: in this situation this will not be possible using the methods described in the non-inferiority appendix.  An approach that nevertheless allows definition and agreement of a clinically important loss of efficacy, without the absolute evidence requirement, would be welcomed as the stated approach is a major barrier to developing alternative, better tolerated agents to substitute other less well tolerated agents	<b>Accepted:</b> In most cases, a substitution design is foreseen. From a regulatory perspective, a non-inferiority design is acceptable and EFS or PFS, as appropriate, are the preferred primary endpoints. It should be recognised, however, that the contribution of a substituted compound to the overall activity of a reference regimen might be hard to define (Main guideline, III.1.4). If the contribution of the substituted compound cannot be directly established based on historical data, also circumstantial evidence indicating that the substituted compound is of clear importance for the overall activity of the reference regimen should be made part of the justification of the selected non-inferiority margin. In addition, the absence of clinically important loss of efficacy of the experimental regimen relative to reference treatments should be demonstrated (Choice of a Non-

		often due to the absence of adequate historical randomised trials.	Inferiority Margin, CPMP/EWP/2158/99). Due to the uncertainties in establishing a proper non-inferiority margin in these cases, it is expected that clinically meaningful reduction in toxicity is demonstrated.
183 – 184	1	<b>Proposed changes:</b> Confounding effects of therapies administered after <u>the</u> end of experimental therapy make <u>endpoints other than</u> DFS more appropriate.	<b>Accepted</b>
184 – 186	1	<b>Proposed changes:</b> This means that CR could be an acceptable primary endpoint <u>when</u> further therapy is scheduled after <u>the</u> end of experimental therapy, <u>such as when</u> induction <u>is</u> followed by consolidation therapy.	<b>Accepted</b>
188 – 189	1	<b>Proposed changes:</b> It is recommended that CR <u>be</u> defined according to....”	<b>Accepted</b>
189	1	<b>Proposed changes:</b> It is recommended that CR is defined according to established clinical criteria, but supportive evidence in terms of MRD as defined by molecular criteria should be sought when applicable. <u>Minimal residual disease (MRD)</u> data, however, should only be used after proven intra- and inter-laboratory validation	<b>Accepted</b>
192	1	<b>Proposed changes:</b> Definition of and/or examples for short- and long term toxicity should be provided	<b>Not accepted.</b> Toxicity depends on mode of action, interference with development and age of the patient. Giving examples is beyond the scope of the guideline.
192	6	<b>Proposed changes:</b> In children reduced or similar toxicity should refer to short time, long time and developmental toxicity, <u>level of education/employment</u>	<b>Agreed.</b> Paragraph will be rephrased to offer more comprehensively potential targets for reducing or showing similar toxicity.  <i>Please note that paediatric sections have been removed from the document and will be included in the upcoming revision of the Addendum on Paediatric Oncology (EMA/CPMP/EWP/569/02).</i>

198	1, 5	<p>The definition of major toxicity should be revisited, basing it purely on SAE differences seems inappropriate. See comment proposed changes to lines 35-39.</p> <p>PFS should also be allowed as an endpoint for superiority studies, with OS/DFS being supportive. Powering studies in DLBCL with DFS as endpoint requires very large sample sizes, possibly delaying significant advances of therapies to be available to patients.</p> <p><b>Proposed changes:</b> <u>In case of severe impairment to patient conditions, e.g GVHD,</u> tThe aim should be to demonstrate increased cure rate or improved survival</p>	<b>Accepted</b>
205 – 206	1	<b>Proposed changes:</b> While “palliation” may be used to cover all situations where “cure” is not the objective, <u>a distinction</u> is made here between...’	<b>Accepted</b>
209 - 211		<b>Proposed changes:</b> Typical conditions include low-grade lymphoma, multiple myeloma, and chronic leukaemias, <u>for which established reference therapies are available and next-line treatment options are likely to be meaningfully active.”</u>	<b>Accepted</b>
209 – 211	1	<p>In these situations alternative endpoints of clinical relevance include such as TFS or time to next treatment, CR and duration of CR, reduction in transfusion requirements if a correlation with prolonged PFS has been previously established.</p> <p><b>Proposed changes:</b> <u>In these situations alternative primary endpoints of clinical relevance include also TFS or time to next treatment, CR and duration of CR, reduction in transfusion requirements if a correlation with prolonged PFS has been previously established. If alternative primary endpoints are used in confirmatory studies agency advice should be considered.</u></p>	<b>Partly accepted:</b> <u>New paragraph:</u> “It is acknowledged that alternative endpoints may be more appropriate in certain situations, such as when maintenance therapy is investigated in areas where this is non-established (see general principles). The wish may also be to enable a long treatment-free interval after intense induction therapy. In these cases and when an until know not established surrogate for long term benefit is planned to be used as primary endpoint, EU scientific advice is recommended prior to the initiation of confirmatory studies.”
209 - 211	5	Alternative endpoints (such as MRD) should be accepted if correlation with prolonged PFS has been previously established.	See above
215 – 218	6	Current wording: Patients withdrawn from therapy prior to	‘

		<p>progression, e.g. due to toxicity, should be followed until disease progression whether on next line therapy or not. This allows for informative alternative PFS analyses including censoring at time of withdrawal, counting as event withdrawal prior to progression and progression off study therapy as event, which in most cases is the preferred primary analysis.</p> <p><b>Proposed changes:</b> <u>The intent-to-treat analysis based on all patients randomized who started the allocated randomized treatment is the preferred primary analysis for the efficacy endpoint (e.g. PFS).</u> Patients withdrawn from therapy prior to progression, e.g. due to toxicity, should be followed until disease progression, whether a next line therapy has been started or not. This allows for alternative informative PFS analyses including: <u>1) censoring at time of withdrawal; 2) counting as event withdrawal prior to progression.</u></p>	<p><b>Accepted.</b> The paragraph, however, has been moved to “Design of confirmatory studies” as being of general relevance.</p>
227	1	<p>Clarification added what is meant by major increase in toxicity since otherwise there could be huge differences in interpretation.</p> <p><b>Proposed changes:</b> <u>In case of irreversible adverse events, severe impairment to patient conditions or otherwise major toxicities,</u> <del>t</del>The principal objective should be to demonstrate <u>prolonged improved</u> survival.</p>	<p><b>Accepted</b></p>
230 – 235	1	<p>In this paragraph it is recommended to include a sufficient number of patients to obtain a precise estimate of possible effects on overall survival. It is not clear how to interpret "precise". Is it recommended to have a proper control of the type I error rate or would some confidence about a trend on overall survival time be sufficient?</p>	<p><b>Comment:</b> The wording was intentionally vague to enable a justification based on the expected benefit in terms of PFS, the increase in toxicity, the expected time period from progression to death, etc. The main message, however, is that it is not appropriate to power the study only in order to meet the primary endpoint.</p>
241	5	<p>In palliative intent studies, BSC or inv choice is acceptable – does this mean, single-arm, non-randomized studies are no longer acceptable in this setting (e.g. recent filing of ofatumumab based on single-arm last-line setting)</p>	<p><b>Comment:</b> Correct.</p>

243	1	<p>The use of superiority with regard to PFS in situations where the reference therapy is known to be active is welcomed. However, in many haematological diseases superiority in terms of ORR/CR (including e.g. durable Complete Response) could also be appropriate. Furthermore, in such situations, where comparisons are made to a proven active comparator, superiority on other measure of clinical benefit (e.g. PFS) suggests actually that a benefit beyond a palliative benefit is required from a regulatory point of view.</p>	<p><b>Comment:</b> The palliative benefit of anti-cancer therapy is thought to be related to anti-tumour activity. In order to estimate the duration of this effect, PFS is considered to be a reasonable measure. In the palliative setting the expected time to tumour progression and death is short. Therefore the need to use alternative endpoints such as ORR is small.</p>
244	1, 5	<p>Better guidelines for acceptable QoL measures and validated scales for studies in which this is appropriate as a primary or secondary measure would be helpful.</p> <p>Double blind design for QoL as mandatory requirement may be difficult to implement in all cases. Therefore it is proposed to have it rather as a recommendation than as a requirement.</p> <p><b>Proposed changes:</b> <u>In the latter case requires it is recommended that the study is conducted under proper double-blind conditions.</u></p>	<p><b>Comment:</b> As for other issues likely to change over time, it is the sponsor's responsibility to select proper instruments for the assessment of QoL. If in doubt, EU scientific advice is a possibility.</p> <p>See below.</p>
243-5, 351-2 and 372-3	4	<p>In three places, the guideline suggests that Quality of Life data "are welcomed" if the study is conducted under double-blind conditions. The concentration on double-blind trials seems overly restrictive. Clearly if the source of the Quality of Life data is the investigator asking the patient questions from a QoL scale then there is an obvious source of assessment bias if the investigator is not blinded. But if the data are provided by the subject at home via an interactive voice response (IVR) system or a handheld electronic diary then there is less opportunity for bias. There are a number of oncology QoL and symptom scales that have been validated for use in electronic applications and, if a double-blind study is not possible or practical, then use of such scales can nevertheless provide valuable data.</p>	<p><b>Comment:</b> While it is possible to "validate" a QoL instrument, it seems rather problematic to "validate" that knowledge of therapy (experimental, active control, or no therapy) would not bias the results. If major differences are shown, the results are probably informative, but in these cases major differences in anti-tumour activity of treatment toxicity are foreseen.</p>
243, etc.	6	<p><b>Comment:</b> it is not at all clear why QoL studies could be considered to produce valid results only in double-blinded</p>	

		<p>settings.</p> <p>If one compares BSC with an experimental treatment (e.g. DNA methylation compound) one can't perform a double-blinded study. Should the results of such a QoL study be <i>a-priori</i> not-valid, despite the fact proper tools have been applied (e.g. QLQ-30), proper design (randomized trial, evaluation at same time points), etc?</p>	See above.
248	1	<b>Proposed changes:</b> .....magnitude of the <u>PFS</u> benefit over the reference regimen..."	<b>Accepted</b>
269	1, 5	<p>The level of BCR-ABL transcript accurately describes the disease's tumor burden as described by e.g. Druker B, et al. Five-year follow-up of patients receiving imatinib for chronic myeloid leukemia. N Engl J Med 2006; 355:2408–2417 therefore this should be acknowledged in the guideline. Also, it was accepted as primary end-point through scientific advice for a few molecules already.</p> <p><b>Proposed changes:</b> Due to too limited experience, <i>BCR-ABL</i> transcript level (“molecular response”) is from a regulatory perspective yet not an acceptable primary outcome measure, but its use as secondary endpoint is non-controversial. <u>However, if new scientific data are available and justify molecular response as primary endpoint, then approval from a regulatory perspective should be acceptable. Scientific advice should be sought to confirm the choice of the endpoint.</u> Technique and response criteria should be justified and comply with updated guidelines and consensus documents.</p>	<b>Accepted</b>
272	1	<b>Proposed changes:</b> Suggest referencing the updated guidelines and consensus documents or providing information on what organizations are working on those.	<b>Comment:</b> There is currently an ongoing international harmonization procedure expected to provide reasoned updates with respect to techniques, etc. Sponsors are advised to follow the literature.
273	1	In order to comply with the definition of the response criteria for CML, the following changes are proposed.	

		<p><b>Proposed changes:</b> <u>In patients with molecular progressive disease, suboptimal response or resistance to</u> a licensed TKI or after secondary efficacy failure, studies may be undertaken in an unselected patient population fulfilling established criteria for non-response or secondary failure, alternatively patients with progressive disease may be enrolled taking into account pattern of mutations if properly justified.</p> <p>Further clarification should be provided what is meant by “unselected population”.</p>	<p><b>Not accepted:</b> The wording primary and secondary resistance was intentionally used to leave open for updated and scientifically and clinically justified definitions of resistance, while the differentiation between “primary” and “secondary” is considered relevant.</p> <p>This referred to the possibility to select or not patients for enrolment based on detected mutations (see rest of the sentence). Wording revised.</p>
273 – 275	1	<p><b>Proposed changes:</b> ...studies may be undertaken in an unselected patient population fulfilling established criteria for non-response or secondary failure; alternatively, patients with progressive disease may be enrolled, taking into <u>account of mutations patterns</u> if properly justified.</p>	<p><b>Accepted</b></p>
279 – 285	1, 3	<p>It is believed that there is no reason for distinction between new TKI and non-TKI. Single arm design should also be relevant for all patients that have no treatment options like TKI resistance (e.g.T315I mutation). Therefore a revised text is suggested.</p> <p><b>Proposed changes:</b> For a new <u>TKI compound</u> and provided that activity in terms of cytogenetic response and duration of response is convincingly high and tolerability and toxicity are well documented and acceptable, single arm studies may still be adequate to support licensure <u>in those patients who have no other treatment options.</u></p>	<p><b>Not accepted:</b> This paragraph should be read in totality where the importance of activity in relation to secondary mutations is emphasised. Based on the experience until now and the benefit – risk of this class of compounds, a mechanistically guided development programme and licensure based on single arm studies is considered to be an option for a third or later line indication.</p> <p>For compound from other pharmacological classes, further justifications are warranted in order to support the use of single arm studies as pivotal for licensure.</p>
288 – 289	1	<p><b>Proposed changes:</b> Signs and symptoms defining intolerance to the prior TKI should be <u>followed, reported on therapy with and distinguished from that related to</u> the experimental</p>	<p><b>Agree, unclear</b></p> <p>Revised wording: Patients intolerant to prior TKI therapy might also be enrolled in these studies, but efficacy should be reported</p>

		compound...”...	separately. Symptoms and signs defining intolerance to the prior TKI should be documented in detail (including grading) prior to inclusion in the study. As class related adverse reactions are common, it is of importance that “cross-intolerance” is excluded as objectively as possible not least due to the subjective nature of “intolerance” in many cases.
290 – 291	1	<b>Proposed changes:</b> ...it is foreseen that add-on studies with a non-TKI <u>that is</u> active in patients with CML will be undertaken.	<b>Accepted</b>
295	1	<b>Proposed changes:</b> Blast Crisis is a very different disease than CML CP and AP and therefore, a separate section for BC is needed as the therapeutical approach is different (chemotherapy 1 <sup>st</sup> line, dasatinib 2 <sup>nd</sup> line and IMP 3 <sup>rd</sup> line). Guidance on study design should be detailed in the guideline, e.g. single arm studies with appropriate historical comparison should be considered.	<b>Partly accepted</b> wording revised: It is foreseen that the vast majority of these patients has been treated with a TKI. For accelerated phase, the guidance given above with respect to patients after failure on a TKI therefore applies. Due to the rarity of blast crisis and the foreseen complexity of the therapeutic situation, EU scientific advice should be considered.
301	1	Depending upon the grade of disease. Intensive chemotherapy and low dose cytarabine are also used among other things. Therefore suggest this line deleted. <b>Proposed changes:</b> <del>Until recently, supportive care or ASCT were the only available treatment options.</del>	<b>Accepted</b> , wording revised
304 and 353 - 358	1	EPOs are currently not indicated for treatment of anaemia in patients with MDS and this guidance should not be seen as endorsing off-label use.	<b>Comment:</b> Nor is intensive chemotherapy or low dose ara-C. This is descriptive, not prescriptive.
308	1	<b>Proposed changes:</b> The clinical course in highly variable and several classification systems have been developed, <u>the main disease ones being the French-American-British and World Health Organisation Classifications, and a prognostic classification called the International Prognostic Scoring System (IPSS).</u>	<b>Not accepted:</b> Current wording sufficiently clear
313 – 319	1, 3	“Recently, new clinical and laboratory variables were identified that might add prognostic information to the IPSS (red blood cell transfusion dependency, high levels of LDH). Sponsors are therefore advised to follow closely the expected refinement of	<b>Not accepted:</b> Current wording sufficiently clear.

		<p>prognostic scores to be used in the design of clinical trials when sufficiently validated.</p> <p>The WHO classification of myeloid neoplasms encompasses disorders that show both dysplastic and proliferative features at time of diagnosis. The following disorders belong to this category: chronic myelomonocytic leukaemia (CMML), atypical chronic myeloid leukemia, juvenile myelomonocytic leukaemia, and myelodysplastic /myeloproliferative disease, unclassifiable (MDS/MPD, U).”</p> <p>Clarify that IPSS classification is acceptable where it is typically the system used by participating physicians.</p> <p><b>Proposed changes:</b> Sponsors are <del>therefore</del> advised to closely follow the expected refinement of prognostic scores <del>to be used</del> in the design of clinical trials. <u>The WHO classification of myeloid neoplasms encompasses both dysplastic and proliferative features, and is when sufficiently validated. However, use of a IPSS classification that is a typical reference for most physicians, is acceptable.</u></p>	
315	6	validated,	<b>Accepted</b>
327	6	Current: as such <b>Proposed change:</b> low risk MDS	<b>Accepted.</b>
320	5	<p>EMA document says: In childhood, MDS should be discerned in adult MDS and paediatric MDS.</p> <p>Please clarify – is this meant to say: In childhood, MDS should be categorized into adult MDS and paediatric MDS?</p>	<p><b>Accepted.</b> Paragraph will be clarified.</p> <p><i>Please note that paediatric sections have been removed from the document and will be included in the upcoming revision of the Addendum on Paediatric Oncology (EMA/CPMP/EWP/569/02).</i></p>
320 - 329	5	Reference to the WHO classification should suffice and would not impact on the guidelines whenever the classification will be changed – which is likely to happen in the near future due to	<b>Comment acknowledged.</b>

		advances in molecular medicine.	
329	1	<b>Proposed changes:</b> Since evolution of bone marrow failure and survival depends....	<b>Accepted</b>
338	1	<p>The statement seems to imply that central reading of <u>all</u> bone marrows and cytogenetics in MDS patients should be done. In case OS is the primary endpoint, baseline BM central assessment (to confirm the diagnosis) is important. However follow-up on bone marrow readings and cytogenetic central readings should not be necessary. In case a claim is based on PFS as primary endpoints, then central readings of cytogenetics or of bone marrow response is necessary.</p> <p><b>Proposed changes:</b> Central read of bone marrows and cytogenetics should be undertaken at least in confirmatory trials <u>where PFS or other efficacy parameters other than OS are used as primary endpoint.</u></p>	<p><b>Not accepted.</b> It is recommended that central reading is undertaken of BM and cytogenetics is undertaken also in studies with survival as primary end point. This is of general relevance and the recommendation is moved to “Design of confirmatory studies”.</p> <p>MDS is a heterogeneous disease where extensive exploratory studies cannot always be undertaken. Thus it is of importance to establish the relationship between BM/cytogenetic response and survival benefit is established as well as possible overall and for relevant subgroups. It is also foreseen that these data would provide supportive evidence of efficacy of importance not least when survival results are not overwhelming.</p> <p>In addition, follow-up of cytogenetics is important from a safety point of view as well, to e.g., correlate modifications with adverse events.</p>
341	6	<p><b>Current:</b> low grade MDS</p> <p><b>Proposed changes:</b> Low risk MDS</p>	<b>Accepted</b>
342 – 346	1	<b>Proposed changes:</b> ... may be too insensitive to capture also relevant differences between treatment groups, <u>especially as</u> transfusion of red blood cells must be individualized due to e.g., concomitant cardiovascular disorders. Loss of need for transfusion for a defined period of time (in combination with improved Hb) is therefore considered <u>an</u> acceptable outcome <u>measure.</u>	<b>Accepted</b>
348 – 350	6	<p>Current: OS and disease evolution must be prospectively assessed to exclude detrimental effects of the test drug that would outweigh documented benefits.</p> <p><b>Comment:</b> We agree. Long term follow-up is sometimes</p>	

		required, to determine, for instance, a possible increase in the risk of second cancer, cardiac toxicity, etc.	
362	1, 5	<p>Low-risk MDS in adults is a disease of the 7<sup>th</sup> decade of life, bordering the median survival of mankind. Although survival data should be collected to evaluate possible detrimental effects due to toxicity, in this era overall survival is not a valid end-point anymore to evaluate the effect of a single intervention.</p> <p><b>Proposed changes:</b> Concerning the respective merits of disease progression-related endpoints <del>and OS</del>, all recommendations expressed.....</p>	<b>Not accepted:</b> This paragraph is meant to cover outcome measures not only for low risk MDS.
p. 8	5	Can QoL be an appropriate primary endpoint with a validated assessment? Particularly in the palliative setting?	<b>Comment:</b> No.
366	1	<b>Proposed changes:</b> ....than for efficacy purposes_(and detection of a clinical benefit).	<b>Accepted.</b>
367 – 368	1, 5	This paragraph emphasizes that confirmatory studies are expected to be randomized and controlled using a <u>licensed</u> medicinal product as reference. The term licensed is considered too strict. The license status of a product may be very different from region to region and would make a global trial difficult to set up.	<b>Accepted.</b> It is the documented benefit/risk that is of importance I do agree. What will be the final sentence then? Could we specify ‘using an active control when available’?
367 – 368	3	<p>For MDS, ORR in a Phase II single arm study should be a sufficient basis for approval.</p> <p><b>Proposed changes:</b> For treatment of conditions such as MDS, available treatment options are limited and may be at best, palliative. Therefore for MDS, a single arm study with an Objective Response Rate (ORR) endpoint should suffice as basis for product approval. If based on an uncontrolled single arm study, the approved indication statement will be qualified since the effect on Overall Survival will be unknown.</p>	<b>Not accepted.</b>
368	1	Leukaemia transformation occurs in 40% of patients, therefore survival should also be presented as supportive.	<b>Accepted.</b> In high risk MDS, survival is the preferred primary end point.

		<p><b>Proposed changes:</b> In principle, PFS (e.g. leukemia-free survival) is an acceptable primary endpoint, <del>but</del> <u>in this case supporting</u> survival data are needed.....</p>	
368	6	<p>Current: PFS (e.g. leukaemia-free survival)</p> <p><b>Comment:</b> according to Cheson et al (Blood 200, Blood 2006), criteria of progression in MDS are broader than the one indicated here, e.g. includes changes in blood and/or bone marrow values as well as regarding transfusion dependency. Leukaemia-free survival (time from the starting point until progression to leukemia or death) may be longer than PFS.</p> <p>Recently (Malcovati et al, JCO, 2007) established the WHO Classification-Based Prognostic Scoring System (WPSS) based on cytogenetic risk, WHO category and transfusion dependency. WPSS was shown to predict survival and leukemia progression at any time during follow-up (<math>P &lt; .001</math>), and its prognostic value was confirmed in the validation cohort.</p>	<p><b>Comment:</b> See above and revised text</p>

## **Proposal to add disease specific information with regard to Chronic Lymphocytic Leukaemia and Non-Hodgkin's lymphoma**

### **4. DISEASE SPECIFIC ISSUES**

#### **Chronic Lymphocytic Leukaemia**

Chronic lymphocytic leukaemia (CLL) is a disease of the elderly, with a median age at first diagnosis of around 65-72 years. The disease generally follows an indolent course and has an overall median survival of about 10 years from first diagnosis. However, the disease has a variable clinical course and the prognosis depends on disease stage and a range of other prognostic factors. Despite high rates of response to treatment initially, remissions are followed inevitably by relapse. Subsequent treatments tend to produce fewer remissions of shorter duration with, ultimately, end-stage, refractory disease. The choice of treatment depends on individual patient and disease-related factors, as well as patient and physician choice.

#### Staging and prognostic factors

The Modified Rai and Binet staging systems are widely used in routine practice and in clinical trials. Both systems are acceptable for use in CLL trials. In large randomized studies, Binet or Rai stage should be considered as a stratification factor and/or an eligibility criterion

A wide range of important biomarkers have been identified including chromosomal aberrations (such as deletions of 17p or 11q [del17p, del11q]), immunoglobulin heavy chain variable region (IgVH) mutational status, and overexpression of CD38 and/or the zeta-chain-associated protein kinase 70 (ZAP-70). Due to the dynamics of the field it is of major importance to follow the evolution with respect to molecular markers and standardisation of molecular techniques used in the assessment of the disease.

#### Treatment intent and primary endpoints

Patients with CLL generally have long duration of survival and improvement in OS primary end point in registration studies might not be a realistic goal. Patients are also likely to have one or more subsequent therapies.. For patients with previously treated or previously untreated advanced/symptomatic disease, progression-free survival is generally an acceptable primary endpoint.. Furthermore, progression-free survival is a rational choice in a disease like CLL which has a long clinical course and in which patients are likely to have one or more subsequent therapies. Progression-free survival should be based on standard criteria - currently NCI-WG criteria are widely used.[Other endpoints e.g. RR, TTF, eradication of minimal residual disease, quality of life etc. could be used as a primary or secondary end points according to several factors e.g. number of prior therapies, presence of subsequent therapies, toxicity of therapy, fitness of the patient etc.

#### Use of reference therapy

The choice of the reference therapy should be assessed taking several factors into consideration such as treatment modalities, e.g. watchful waiting, available established therapy, previous treatment received and patient general condition.

## **Non-Hodgkin's Lymphoma**

Non-Hodgkin's lymphoma (NHL) is a group of haematological malignancies which are primarily affecting lymph nodes but can also involve bone marrow, spleen, liver, and other extranodal and extralymphatic organs. Whereas the malignant cells are always of lymphoid origin, NHLs are heterogeneous in terms of pathophysiological, histological and clinical characteristics. The vast majority (>90 %) of NHLs are B-cell malignancies and only a minority of cases resemble T- or NK-cell neoplasms.

From a clinical and therapeutic perspective, an important classification is aggressive NHL (mostly diffuse large B-cell lymphoma, DLBCL, and mantle-cell lymphoma, MCL) vs. indolent NHL (mostly follicular lymphoma, FL). Whereas aggressive NHL can potentially be cured with the current treatment regimens, indolent NHL usually relapses and requires repeated treatment

In some disease entities, the disease risk is a key factor for treatment selection. Up to now, the International Prognostic Index (IPI) which includes 5 risk factors (age, disease stage, extranodal involvement, LDH elevation, performance status) has been commonly used to describe the disease risk. Recently, a modified risk model was developed for patients younger than 60 years (age-adjusted IPI, aaIPI) which includes only 3 factors (disease stage, LDH elevation, performance status). In addition, more entity-specific risk models have been described for follicular lymphoma (FLIPI) and mantle-cell lymphoma (MIPI). As the disease risk assessment models are becoming increasingly relevant for the description of the patient populations which are included in clinical studies and for the selection of the type and intensity of treatment, the evolution in risk models should be carefully followed and considered where appropriate for clinical studies. Importantly, similar variations are present for the definition of bulky disease for which different thresholds for the minimal bulk diameter have been defined. Therefore, also for the definition of bulky disease, the selection of the appropriate cut-off diameter is important to sufficiently describe the study population.

### **Aggressive NHL**

In aggressive lymphoma, CR and CRu are usually considered sufficient responses to therapy depending on the response assessment guidelines. As therefore active or non-active anatomical residues are present even after successful treatment, PFS is considered the appropriate survival-related endpoint. However, as in particular for younger patients with a higher disease risk, high-dose treatments with subsequent HSCT are a treatment option with clear curative intent, DFS may be preferred for example for high-dose treatment schedules.

### **Indolent NHL**

In indolent lymphoma, PFS is considered as the preferred primary study endpoint due to the frequent presence of lymphoma residues after successful treatment.

#### **Response assessment**

The response assessment is described in guidelines of the NCI-sponsored International Working Group. Deviation of the response assessment (e.g. not using assessment techniques like PET) from the recommended NCI-sponsored International Working Group should be justified.