7 December 2020
EMA/658070/2020

Overview of comments received on ICH guideline E14/S7B on clinical and nonclinical evaluation of QT/QTc interval prolongation and proarrhythmic potential - questions & answers (EMA/CHMP/ICH/415588/2020)

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

Stakeholder no.	Name of organisation or individual
1	Peter Ravn Brinck, Novo Nordisk A/S
2	Voisin Consulting Life Sciences
3	Boehringer Ingelheim Pharma GmbH & Co. KG
4	AstraZeneca
5	Borje Darpo MD PhD, Associate Professor of Cardiology (SE)
6	Georg Ferber PhD, Statistik Georg Ferber GmbH (CH)
7	Simon Hebeisen PhD, B'SYS GmbH (CH)
8	Helen Prior PhD, DSP, ERT, National Centre for the Replacement Refinement & Reduction of Animals in Research (NC3Rs)(UK)
9	EFPIA
10	CSL Behring GmbH

Please note that comments will be sent to the **ICH E14/S7B IWG** for consideration in the context of Step 3 of the ICH process.

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

1. General comments – overview

Stakeholder no.	General comment (if any)
4 and 9	General comment: *Either* merge revised 5.1 and 6.1 into a single question *or* focus 5.1 on the requirements for clinical concentration-QTc modelling, and move all discussion of the role of the integrated nonclinical-clinical risk assessment into 6.1. Rationale for proposed change: Q6.1 describes a route to integrate non-clinical and clinical QTc data in specific scenarios when it is not possible to perform a conventional thorough QT study (typically taken to refer to oncology therapies). With the acceptance of concentration-QTc modelling (as described in Q5.1), and the proposed integration of non-clinical data into this evaluation when it is not possible to sufficiently exceed the maximum clinical exposure, the rationale for separating these 2 scenarios has decreased. Question 5.1 now includes a short section on the integration of non-clinical QTc data into the risk assessment which introduces the potential for confusion: "are we in a 5.1 or 6.1 scenario?" Approximately 80% of all thorough QT studies are negative (Wisniowska et al 2020), illustrating the limited value derived for patient safety from this evaluation. By extending the use of the 6.1 scenario to any new compound where the combined nonclinical and clinical QT assessment indicates low risk (i.e. not just for oncology therapeutics), greater value would be derived for drug development. Merging the revised elements of 5.1 and 6.1 into a single question, or focusing 5.1 simply on clinical concentration-QTc and 6.1 on the role of the integrated non-clinical-clinical assessment, would support this aim. General comment: In clinical oncology often Q&A 6.1 and Q&A 5.1 are applicable. However, from experience, it remains challenging to navigate through these recommendations in many oncology small molecule programs where trials are in patients and data are noisy (no time-matching, comorbidities, and comedication). It would be helpful if the WG could formulate a more comprehensive recommendation under these scenarios. Additionally, medium QT signals >5ms <15ms are not uncommon for oncology drugs, yet the relevance of these signals against benefit-risk is often limited. Further formalization of distinguishing 'no QT effect' from 'no large QT effect' or other intermediate wording regarding QT risk would be helpful in oncology setting, where establishing 'no QT effect' is a nice-to-have, but rarely relevant. Can this be clarified?
6	The comments below refer to statistical aspects to the document.
8	There is currently no reference to the 3Rs (reduction, refinement, replacement) within the original ICHS7A and ICHS7B guidelines nor the proposed Q&As. To reflect the importance of these principles globally, and to harmonise with other recent ICH guideline updates and/or new guidelines, can an addition to section 1.4 General Principles be considered, similar to the following 'Appropriate efforts should be considered to continue progress in the 3Rs of Reduction, Refinement, and Replacement in the use of animals'.
9	We are generally aligned with the intent of the concept paper and the new ICH E14/S7B Q&As. We welcome the opportunity for the non-clinical work to impact the clinical implementation and interpretation of the clinical cardiac repolarization assessment. This has been the standard practice for the industry sponsors for two

Stakeholder no.	General comment (if any)
	decades and the low positive TQT rate described in the ICH public webinar on the new Q&As would support the success of this practice. Aligning the regulatory assessment to this totality of evidence approach is a valuable step. We suggest that continued vigilance on the question of specificity of the assessment is very important. It is acknowledged that the use of a small QTc prolongation is a sensitive, but not specific biomarker for the arrhythmia torsade de pointes. The Stage 2 Q&As are clearly aligned to the specificity question with an explicit intent to evaluate the proarrhythmic potential of compounds which prolong the QTc interval. However, there may be a danger that the manner in which the Stage 1 Q&As are implemented could actually increase the sensitivity and decrease the specificity of the assessment even further. This would create a significant false positive rate given that the pretest likelihood of having a proarrhythmic drug is diminished, owing to greater awareness of structural properties favouring hERG block, improved chemical matter in general, and the increased prevalence of non-small molecule modalities in drug development. The 'false positive' issue would likely manifest in the preclinical space, prior to any potentially constructive regulatory dialogue, and impact the number of molecules coming forward for clinical development or increase the time to 'discover' an apparently appropriate compound. Q&A 1 and Q&As 5.1 and 6.1 *Areas of the draft Q&A where greater clarity is needed* There is flexibility in defining the margin to be used for the hERG assessment (relevant to Q&A 1 for S7B and Q&As 5.1 and 6.1 for E14) and in how a positive effect and 'similar sensitivity' is defined for the *in vivo* QTc assessment (ICH S7B Q&A 3). Flexibility is good in general, but it leaves uncertainty in how to implement the Q&As. It also leaves the possibility of different regional interpretation open. The latter would seem counter to the goals of ICH. For better or worse ICH E14 was very explicit about the expectation and interpretation around 10ms QTc prolongation. The hERG margin has three dimensions, all changing with the new Q&As: a standard hERG protocol is suggested along with technical study recommendations, the plasma protein binding will be capped at 99% for highly bound compounds, and we heard in the public webinar that the denominator clinical exposure would be the "high clinical scenario". In principle all these are reasonable, but we've never seen all of these combined for reference agents with or without known risk of torsade de pointes. This obviously means that a specific recommendation cannot be made, but uncertainty will exist until this type of data representation is available and the interpretation of acceptable margin has been discussed. In the public webinar on the ICH Q&As a figure suggested that 100% or near 100% sensitivity in the margin threshold used would be expected. This level of sensitivity would likely come at the expense of specificity; this may be counter-productive in the overall assessment of cardiac repolarization effects. Specific recommendations would be that the margin of 30-fold has served us well since it was first proposed in 2002; we would expect a margin of ≤30-fold to be the likely margin chosen. The comparator compounds to

Stakeholder no.	General comment (if any)
	define and demonstrate should be known to be associated with torsade under normal therapeutic use and not from the "conditional risk" or "possible risk" lists since it is highly unlikely these are associated with torsade under normal therapeutic use and a margin using a therapeutic concentration as the denominator will be too large. Even for compounds known to be associated with torsade de pointes the 'high clinical scenario' is more likely to be relevant than the therapeutic concentration. A list of compounds with the 'high clinical scenario' exposures and agreed plasma protein-binding fraction should be provided to sponsors. Similarly, how the *in vivo* assay is 'powered' and given an optimal cut-off is a permutation of magnitude of QTc change and concentration-multiple which has never been fully discussed. We feel confident that we achieve adequate levels of sensitivity already with our current study designs (n=8 double Latin Square cross-over). We acknowledge that this could be demonstrated more clearly in the report text for the less familiar reader. We also acknowledge that there is general concern that more stringent sensitivity requirements may increase the number of animals necessary for many sponsors unless the emerging concentration-QTc assessment proves to have adequate performance characteristics. There is an 'information void' for many sponsors in this regard. This seems to be a lot of uncertainty for a guideline close to finalization. In E14 Q6.1 there is discussion of an imbalance in events in the safety database. The choice of the specific events needs careful consideration as some cardiovascular events may not be clearly associated with cardiac repolarization and also occur at rates so low that the opportunity for random imbalance would be significant. This could again impact the performance characteristics – sensitivity and specificity – of this type of assessment. *Areas where your company is supportive of the proposed approach(es)* We are very supportive of the concept of using the nonclinical data alongside clinical data in a totality of evidence approach to proarrhythmia assessment and interpretation. We'd encourage its use in the widest possible range of scenarios. The answers given in the public webinar were very encouraging in this regard. It seems the new approach of using the nonclinical "double negative" in conjunction with more limited clinical data could be leveraged to arrive at a 'no QTc effect' interpretation, for the purposes of late stage clinical monitoring, and 'low risk' label. This E14 Q&A 6.1 had been largely interpreted as applying to oncology agents, but the verbal answers suggested it could apply to many more molecules provided they met the "double negative" standard. Molecules with low systemic exposures and an increased range of non-small molecule modalities (e.g. peptides and oligonucleotides) would be obvious inclusions. *Areas of concern with the proposed approach(es)* As stated above we are aligned with the approach being proposed, our concerns would be around any unintended consequences in how it is implemented globally. If we are also looking for an approach which is no more sensitive than the current approach, but with more specificity as has often been articulated in publications and

Stakeholder no.	General comment (if any)
	meetings some of the missing detail around margin and *in vivo* sensitivity will become very important. We note that the original concentration-QTc analysis in the clinic was very encouraging but then it was coupled to a need to exceed even the "high clinical scenario" exposures in order to reassure on assay sensitivity. This limited the utility of concentration-QTc as a TQT substitute. That limitation is now being addressed by the new Q&As, but the potential will only be fully realized if the details of hERG and *in vivo* margin are set appropriately. We have concerns that the flexibility and existing uncertainty described above leads to different regional standards in interpretation. We have concerns that the standards being described by the S7B Q&As are portrayed as higher and 'best practice'. It would be unfortunate if these gave the impression that many sponsors were not already conducting studies to a high standard and that an even higher standard became the expectation. If this were then to become the expectation for all molecules prior to FHD that would be inappropriate. There is no explicit intent or need in the Q&As to have that occur.

2. Specific comments on text

E14 Q&A

Question no.	Stakeholder no.	Comment and rationale; proposed changes
5.1	1	**Comments:** The current paradigm seems to suffer fewer limitations when the drug in question can interact directly with ion channels proportional to drug concentration and without delay. For drugs that do not interact with ion channels the QT-prolongation assessment appear less relevant and a discussion of different drug classes e.g. highly selective drugs like peptides that may not require a TQT study would be welcome. For these compounds collecting high quality ECG at time points that may not be well defined yet in early human studies is an unnecessary burden. Non-clinically the double negative approach is challenged by the lack of scientific rationale for performing hERG test of selective peptide drugs (target specificity, size and minimal presence in cytosol at inner cavity of the hERG channel). Furthermore, many peptides have long acting profiles which does not support cross-over designs in non-rodent cardiovascular *in vivo* studies resulting in parallel study designs with reduced sensitivity and the potential need to increase group size substantially using more animals including non-human primates. To align with ICH S7B Q&A #3.2, it is suggested to clarify in ICH E14 Q&A #5.1 that the nonclinical *in vivo* assay should be conducted at exposures which cover the anticipated high clinical exposure scenario. The wording "no QTc prolongation in an *in vivo* assay" is unclear and could mean any statistically significant effect related to treatment OR an effect above a given threshold in each species. This should be clarified. While the expectation to the sensitivity in the nonclinical *in vivo* assay is clarified when used to support an integrated clinical and nonclinical risk assessment as described in ICH E14 Q&A 6.1, the same is not the case for ICH E14 Q&A 5.1. Could it be elaborated in which range the minimum detectable difference should be in order to be an assay of "sufficient sensitivity". The risk here is to not meet expectations or to use an increasing number of animals including nonhuman primates to achieve a higher than expected sensitivity.
5.1	3	**Comments:** Please consider to include a hint for QTc assessment for drugs with heart rate effects.
5.1	3	**Comments:** "sufficient sensitivity" of the in vivo assay…Can this be specified?
5.1	3	**Comments:**

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		"exceed clinical exposures"…How many multiples are expected here? Is the worst clinical scenario x2 meant here as per guideline?
5.1	3	**Comments:** "[…] interval for the QTc effect of a drug treatment as estimated […]" and "[…] should be <10 ms […]" Is "mean QTc" meant here? I yes, please specify.
5.1	4 and 9	**Comments:** The definition of "exposures…that exceed clinical exposures" should be clarified. **Rationale for proposed change:** Unless there is a clear definition of by how much the exposures in the non-clinical in vivo QT assay should exceed the clinical exposure, regulators from different regions may interpret this differently resulting in sponsors being able to use non-clinical QTc data to support a clinical concentration-response analysis in one region but not another.
5.1	4 and 9	**Comments:** "Sufficient sensitivity" should be defined **Rationale for proposed change:** The proposed wording supports use of a high-quality nonclinical in vivo QT assay following the principles outlined in Q3.1-3.5 to supplement the clinical QT evaluation whilst providing flexibility in how the assay should be performed. Since there is no current consensus on the technical details of an in vivo QT assay that meets Q3.1-3.5, the proposed wording makes it clear that by satisfying these requirements the assay would be considered to have acceptable sensitivity. **Proposed change:** (2) no QTc prolongation in an in vivo assay with defined sensitivity conducted at exposures of parent compound and human-specific major metabolites that exceed clinical exposures, such that QTc prolongation at sufficiently high clinical exposure multiples can reasonably be excluded.
5.1	4	**Comments:** Would it be sufficient to say "above" rather than well above; "well" sounds like a value word and does not give more clarity than above when not specifying a minimum. Please clarify.

5.1	5	**Comments:** Under Important considerations, point 4), it is stated that a separate positive control would not be necessary if either of the following is met: *There are data characterizing the response at a sufficiently high multiple of the clinically relevant exposure (see ICH E14 Section 2.2.2);* *"If the maximum therapeutic exposure has been fully covered in the clinical ECG assessment (e.g., concentrations representative of the maximum recommended dose at steady-state in situations of intrinsic and/or extrinsic factors that increase bioavailability), but sufficiently high multiples* ***cannot be obtained (e.g., for reasons of safety, tolerability, saturating absorption),*** *then a nonclinical integrated risk assessment that includes the hERG assay, an in vivo QT assay, and any follow up studies can be used as supplementary evidence. See ICH S7B Q&A 1.1 for details; in summary, the nonclinical studies should include (1) a hERG safety margin higher than the safety margins computed under the same experimental protocol for a series of drugs known to cause torsade de pointes (TdP) and (2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and human-specific major metabolites that exceed clinical exposures."* Current FDA practice, not necessarily followed by all other regulators, has been to request a TQT study with a positive control in cases where very high concentrations cannot or has not been achieved in e.g., the First-in-Human (FIH) study. The proposed change will therefore, to some extent, lower this requirement, i.e., decrease the number of TQT studies, and enable acceptance of robust high-quality ECG data, using C-QTc analysis, supplemented by non-clinical data, to demonstrate that the drug does not cause clinically relevant QT prolongation. It should then be emphasized that as the text is written, this applies only to drugs for which sufficiently high concentrations cannot be obtained (e.g., for reasons of safety, tolerability, saturating absorption). If the assumption that '*the maximum therapeutic exposure has been fully covered'* is shown to be correct, higher concentrations may, in fact not be seen in patients, including those with impaired clearance of the drug and those at risk for proarrhythmic events, and the revised text then gives a path forward without performing a stand-alone TQT study. In this context, it is important to point out the role of high concentrations in terms of detecting the QT effect of a drug (1). We know that high concentrations are key for the ability of C-QTc analysis to detect small QTc effects, and thereby increase our confidence in the data. In the example shown below, the QT effect of a drug became apparent only when a higher dose group was added to the analysis in a multiple ascending dose (MAD) study (2). The graph shows that for multiple doses from 5 mg to 90 mg, the C-QTc relationship was shallow (blue lines) and an effect on the placebo-corrected, change-from-baseline QTcF ($\Delta\Delta$QTcF) of more than 10 ms could be excluded throughout the observed range of concentrations. When adding a higher dose, 150 mg, the C-QTc relationship became positive (red lines)

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		and similar to the relationship seen in a TQT study, which also included the 150 mg dose. This is also an experience that I, and other colleagues involved in this area, have encountered in other studies. It seems prudent to underline that 5.1 is restricted to those cases where sufficiently high concentrations cannot be obtained, rather than broadly applying these criteria. I therefore suggest that the current wording under b) should be revised to include a statement along the following lines: *If sufficiently high multiples of maximum therapeutic exposure were not obtained in the clinical ECG assessment, and reference is made to 6.1, it should be clarified why sufficiently high exposures cannot be achieved.* **Please see also Appendix 1**
5.1	6	**Comments:** In the sentence: *However, hypothesis testing based on a by-time point analysis (intersection-union test or point estimate and confidence intervals) is inappropriate in studies designed for a concentration-response analysis, if not powered to assess the magnitude of QT prolongation for each time point.* **Proposed change:** I would suggest to add "and dose group of interest", since the by timepoint analysis needs to be performed within a dose group.
5.1	9	**Comments:** Does A more detailed explanation for the analysis require a pre-specified model inclusion of intrinsic and extrinsic covariates in the analysis plan as described by concentration will be helpful. Including some of the white paper by Garnett et al? Or is published work in the C-R analysis seen as model building without a priori model specified? Reference section will be beneficial.
5.1	9	**Comments:** "Concentration-response analysis, in which all relevant data across all doses are used to characterize the potential for a drug to influence QTc" , relevant doses, very low subtherapeutic doses will not be informative. Can it be more specific, eg data that is several fold below therapeutic dose and above should be included. **Comments:** Item # 3 (Page 7): If the plan is to pool the data from several studies for concentration-response modelling, at what stage should the effort be

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		finalized? With the first study to be pooled or after a few of the studies are completed? When should a SAP be prepared?
5.1	9	**Comments:** Should extrapolation outside the range of studied concentrations be discouraged?
5.1	9	**Comments:** Item 3 talks about testing for heterogeneity when it is desired to pool data from multiple studies. It would be helpful to provide some criteria/guidance for assessing poolability; e.g., p-value for the study effect to reject poolability or which test to use etc.? This will help with pre specification in statistical analysis plan. **Proposed change:** Please provide some criteria for assessing poolability of data from multiple studies.
5.1	9	**Comments:** It would perhaps be useful to include more guidance on what is meant by high quality ECGs in the context of early phase studies. **Proposed change:** Please add criteria for "high quality" ECGs.
5.1	9	" 1. Data can be acquired from first-in-human studies, multiple-ascending dose studies or other studies provided that the concentrations achieved are well above the exposure at the maximum therapeutic dose at steady-state, and reflect high exposure scenario situations…" **Comments:** As per the EMA "Guideline on strategies to identify and mitigate risks for first-in-human and early clinical trials with investigational medicinal products" (2018), section 7.5 specifies that "In general, the maximum exposure of healthy volunteers should be within the estimated human pharmacodynamic dose range." and that "A trial design using a MTD approach is considered to be inappropriate for healthy volunteers." Is the high exposure scenario in adequacy with the pharmacodynamic dose range?
5.1	9	**Comments:** Questions on the Q&A: We would like some clarifications in case of active metabolite(s) contributing to the PD effect and to the safety in addition to the active parent compound (PD and safety), should there be a "parent conc-ECG analysis" + "active metabolite(s) conc-ECG analysis" as well as an "integrated" conc-ECG

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		analysis? What kind of modelling approach would be recommended? As Parent and active metabolite(s) pharmacokinetic profile may be different.
5.1	9	**Comments:** "If there is an intention to pool data from multiple studies, it is important to test for heterogeneity." We would like to know if it means that a statistical test is required, or would a graphical check could be enough in this particular case ?
5.1	9	**Comments:** Paragraph starting by "Concentration-response analysis……." While this is a great scientific discussion, it reads like a research manuscript and it is not clear what the actual answers are. There are 2 questions - need to answer each one succinctly and directly. Paragraph starting by "If the maximum therapeutic exposure has been fully covered in the clinical ECG assessment (e.g., concentrations representative of the maximum recommended dose at steady-state in situations of intrinsic and/or extrinsic factors that increase bioavailability),…. " How is this defined? Need to provide specific guidance. Section starting by "other uses"; This section does not provide answers only comments. Consider its value as part of the guidance document.
5.1 and 6.1	9	**Comments:** The distinction between E14 Q&A 5.1 and 6.1 is subtle. It would seem that either the necessary criteria for 5.1 should be less stringent e.g. based on the hERG margin and not mentioning the *in vivo* QTc, or Q&A 5.1. and 6.1 could be merged. Q&A 6.1 was largely positioned as an oncology scenario; however it could be positioned that the presence of a nonclinical "double negative" and a modest amount of clinical data is sufficient to describe any compound as low risk for proarrhythmia. This would place an emphasis in the document on the integrated risk assessment Q&A and actually move it from an ICH S7B Q&A to an ICHE14 and S7B Q&A. This is perhaps signalling that both documents should effectively be revised rather than supplemented with an increasing number of Q&As then revisions on Q&As.
5.1 and 6.1	9	"…(2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and human-specific major metabolites that exceed clinical exposures." **Comments:**

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		This statement might lead to "false positive" in vivo data, e.g. in cases where a QTc prolongation was seen at 30x multiples to the human therapeutic exposure, but not any more at 3x or 10x multiples. Following the wording in this section, a QT prolongation at ANY multiples of clinical exposure would render an in vivo study positive. The same statement is true for page 9, Q 6.1. 1. and page 11, Q1.1. **Proposed change:** "…(2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and human-specific major metabolites that exceed clinical exposures **within a relevant range**." *OR* "….(2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and human-specific major metabolites that **cover the anticipated high clinical exposure scenario**.
5.1 and 6.1	9	**Comments:** The integrated nonclinical and clinical QT/QTc risk assessment should include: 1. The hERG assay, an in vivo QT assay, and any follow-up nonclinical studies, especially those selected to overcome the challenges encountered in the clinical studies (see ICH S7B Q&As 1.1 and 1.2); and 2. Alternative QT clinical study designs incorporating ECG assessments with as many of the usual "thorough QT/QTc" design features as possible (see ICH E14 Section 2.2 and Q&A 5.1). Follow-up non-clinical studies are to be included into the integrated risk assessment based on the above. However, the decision-making about low likelihood of proarrhythmic effects due to delayed repolarization only includes the hERG safety margin, which in its turn requires considering effects on other cardiac ion channels (page 10 Q1.1). Should, therefore, testing compounds effects on additional ion channel or cardiomyocytes be considered for the evaluation of the non-clinical studies in general? **Proposed change:** The nonclinical studies should include (1) a hERG safety margin higher than the safety margins and computed under the same experimental protocol for a series of drugs known to cause torsade de pointes (TdP) ***considering compound effects on other cardiac ion channels*** and (2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and human-specific major metabolites that exceed clinical exposures.

Question no.	Stakeholder no.	Comment and rationale; proposed changes
5.1 and 6.1	9	**Comments:** The definition of "exposures…that exceed clinical exposures" should be clarified. **Rationale for proposed change:** Unless there is a clear definition of by how much the exposures in the non-clinical in vivo QT assay should exceed the clinical exposure, regulators from different regions may interpret this differently resulting in sponsors being able to use non-clinical QTc data to support a clinical concentration-response analysis in one region but not another.
6.1	1	**Comments:** The wording "no QTc prolongation in an in vivo assay" is unclear and could mean any statistically significant effect related to treatment OR an effect above a given threshold in each species. This should be clarified. To align with ICH S7B Q&A #3.2, it is suggested to clarify in ICH E14 Q&A #6.1 that the nonclinical in vivo assay should be conducted at exposures which cover the anticipated high clinical exposure scenario.
6.1	2	**Comments:** page 8, ICH E14, #6.1, Section Decision-Making, Nonclinical studies: We suggest mentioning the quantitative expected ratios to demonstrate a low risk in the hERG assay and a negative nonclinical QTc evaluation -a low risk in the hERG assay is defined as >30-fold ratio between the hERG-IC50 value and the clinical free therapeutic drug level -a negative nonclinical QTc evaluation is defined as >10-fold ratio between the highest free plasma drug level with no QTc effect and the relevant clinical free therapeutic drug level **Proposed change:** The nonclinical studies, following best practice considerations for in vitro studies (see ICH S7B Q&A 2) and in vivo studies (see ICH S7B Q&A 3), show low risk which includes (1) a hERG safety margin ~~higher than the safety margins computed under the same experimental protocol for a series of drugs known to cause TdP~~ **>30-fold ratio between the hERG-IC50 value and the clinical free therapeutic drug level**; and (2) no QTc prolongation in an in vivo assay of sufficient power to detect a QTc prolongation effect of a magnitude similar to dedicated clinical QT studies and at exposures of parent compound and human-specific major metabolites that exceed clinical exposures, **considering that a negative nonclinical QTc evaluation is defined as >10-fold ratio between the highest free plasma drug level with no QTc effect and the relevant clinical free therapeutic drug level** (see ICH S7B Q&A 1.1 for details)

Question no.	Stakeholder no.	Comment and rationale; proposed changes
6.1	3	**Comments:** "[…] comparison is not possible; safety considerations […]" **Proposed change:** Please change semicolon to comma for better understanding. Or, if the following text is meant as listing, please consider to change to colon
6.1	3	**Comments:** "to overcome the challenges"… If the clinical QT assessment has to be done in patients on concomitant therapy (if tolerability does not allow to go into healthy subjects), how should the nonclinical study look like to overcome this challenge?
6.1	3	**Comments:** "(i.e., <20 bpm)"… Why is the threshold for confounding heart rate effect now defined as 20 bpm (used to be 10 bpm in the white paper)? Is "confounding heart rate effects" different from "clinical significant heart rate effect" in the sense of this guideline vs. the white paper? **Proposed change:** It should be made clear that "mean" heart rate effects are meant. Moreover, instead of "…confounding heart rate effects…", "confounding heart rate changes…" should be use.
6.1	3	**Comments:** "Advanced methodologies for controlling or correcting heart rate changes"…Is QTcI meant here? Which other advanced methodologies for correcting heart rate would apply?
6.1	3	**Comments:** "defined as ΔQTc greater" **Proposed change:** Please include "mean" and shouldn't this be "mean ΔΔQTc" since it is the placebo-corrected change from baseline?
6.1	3	**Comments:** What is meant with "effect on ΔQTc"
6.1	3	**Comments:** "as large as 20 ms"…What is the rationale for "20 ms"?
6.1	4	**Comments:**

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		Is it possible to detail when hERG / in vivo studies would not be performed, or are not considered a prerequisite? e.g. ICH S9? **Rationale for proposed change:** Alignment to other ICH guidelines **Proposed change:** If nonclinical studies do not show low risk (or are not performed i.e. when following ICHS9).
6.1	4	**Comments:** The bullet 3 needs to be clarified for the requirement as pre-approval & post approval. Referring to the available cardiovascular safety data from phase 1 to 3 (pre-approval) & including post marketing reported CV adverse events (post approval)? Please clarify.
6.1	4	**Comments:** Could you please clarify how the AEs should be presented in the database: Per study? as there will be studies with different durations, different doses/exposures, different patient populations (inclusion/exclusion criteria and as such cardiac comorbidities will vary), monotherapy versus combinations of drugs? Healthy volunteers (if applicable) and patients with target disease separately? Number of patients/HV per Type of AE relative to total number of patients exposed? All CV SAE/AEs or just those suggestive of arrhythmia/QT prolongation? Please clarify.
6.1	4	**Comments:** Can you provide clarity on the impact/weight of each of the 3 components? Component 1 and 2 are quantitative and if they do not suggest QT prolongation; component 3, which is more difficult to quantify adequately, due to data heterogeneity and impact of comedications with possible prolonging effect, etc, should carry less weight. In the situation where the non-clinical assessment and clinical ECG assessment are clearly negative, a numerical imbalance in the AE database would not be considered to be a equivalent to either a positive non-clinical or clinical QTc assessment, nor a proarrhythmic signal. This should be clarified in the Q&A.
6.1	4	**Comments:** Please clarify point (2) in the decision making section. What would enable a low risk conclusion if points 1 and 3 are achieved? Should the lack of clinical relevant QTc prolongation be defined as dQTc <10ms with a upper bound of the CI < 20 ms?

Question no.	Stakeholder no.	Comment and rationale; proposed changes
6.1	4 and 9	**Comments:** Confounding HR effect considered if HR change is >20 bpm. If this is the defined threshold to apply advanced methodologies for controlling or correcting for heart rate changes, it would be further helpful to specify such methods, please clarify.
6.1	4 and 9	**Comments:** The WG proposes that the strength of the clinical ECG data depends on the upper bound of the two-sided 90% confidence interval around the mean ΔQTc estimate. For the purposes of sizing studies, ensuring statistical power and pre-specification thereof, it would be helpful to have more guidance on this topic. Could the WG provide guidance on the impact of various dQTc thresholds or confidence interval widths on labelling and clinical program? **Rationale for proposed change:** This would allow sponsors to size and design their studies and analyses. Also, it would allow them to determine the strength of evidence needed against the benefit-risk profile and clinical relevance for the pursued indications. Please clarify.
6.1	4 and 9	**Comments:** Sentences unclear, please clarify. **Proposed change:** An integrated nonclinical and clinical QT/QTc risk assessment can be particularly valuable under scenarios where a placebo-controlled comparison is not possible such as when safety considerations preclude administering supratherapeutic doses to obtain high clinical **exposures** and/or safety or tolerability prohibit the use of the product in healthy participants.
6.1	5	*In situations where it is not possible to evaluate the QT/QTc effects at higher exposures than are anticipated with the recommended therapeutic dose, it is particularly important that the nonclinical in vivo studies are conducted at exposures exceeding the clinical therapeutic exposures.* **Comments:** An integrated QT/QTc risk assessment can also be particularly valuable for drugs with confounding heart rate effects (i.e., >20 bpm) that could impact accurate determination of the QTc…
6.1	5	*A totality of evidence argument based on the results of an integrated nonclinical and clinical QT/QTc assessment could be made at the time of marketing application. To support a drug as having low likelihood of proarrhythmic effects due to delayed repolarization, the assessment should demonstrate the following:*

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		2. *The high-quality ECG data (see ICH E14 and E14 Q&A 1) collected in the alternative QT clinical assessment do not suggest QT prolongation, generally defined as ΔQTc greater than 10 ms, as computed by the concentration-response analysis (see E14 Q&A 5.1 for details) or the intersection-union test. The strength of the clinical ECG data depends on the upper bound of the two-sided 90% confidence interval around the mean ΔQTc estimate…* **Comments:** **'*Low likelihood of proarrhythmic effects*':** Since this refers to a claim that a sponsor can make at the time of marketing applicant, it should be noted that the risk/benefit assessment and labeling are performed separately by each regulatory authority (see E14 5.2), and may therefore vary across regions, especially when an effect on ΔQTc > 10 ms cannot be excluded. It may well be that regulators will see this in a similar way for a drug with a small effect (e.g., 4 ms with an upper bound of the 90% confidence interval (UB90%CI) of 12 ms, as in Dr. Garnett's example, slide 18 and 19). It is, however, not evident, in my view, that the same is true for a drug with a larger effect, still within the 6.1 definition, e.g., mean ΔQTc of 9 ms, UB90%CI of 17 ms. If a harmonized regulatory approach is desired, it seems better to retain the threshold that most parties can agree on, i.e., exclusion of a 10 ms effect (i.e., UB90%CI < 10 ms). The issue I see with Q&A 6.1, is that the consequences for patient studies of '*low likelihood of proarrhythmic effects'* are not described. Even though it is clearly stated that the claim about 'low proarrhythmic effect' can be made at the time for marketing application, the clinical QT evaluation will in many cases be performed before pivotal studies are initiated, e.g., by applying C-QTc analysis on data from the FIH study in cancer patients. It can then be argued that a drug that can be categorized as having *low likelihood of proarrhythmic effects* based on 6.1 criteria, can be given safely to patients in Phase 3 trials, without exclusion criteria or cautionary statements in regard to concomitant medications with drugs that are known to cause QT prolongation, or to patients at risk based on e.g., family history of LQTS, cardiovascular disease or hypokalemia and without ECG monitoring. As defined under 6.1, a drug that causes a mean ΔQTcF of 9 ms with an UB of 17 ms can be viewed as 'safe' from this perspective. I disagree that such drug can be taken into large patient trials without specified exclusion criteria and precautions and without ECG monitoring, but much more importantly – I do not believe there is consensus across regulators on this point. This means that sponsors will not know what to expect and that the desired harmonization across regulators is not achieved. If a new term is introduced into the guidance, I think it is important to describe the consequences for subsequent patient studies by referring to the *E14 section 2.3*. I suggest that it would be better to keep the threshold on which there is consensus (exclusion of a 10 ms effect) and then allow

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		regulators to make case-by-case decisions, depending on the severity of the indication and the unmet medical need. Alternatively, the text under 2) can be revised along the following lines (added text in bold): *….generally defined as ΔQTc greater than 10 ms, as computed by the concentration-response analysis (see E14 Q&A 5.1 for details) or the intersection-union test. The strength of the clinical ECG data depends on the upper bound of the two-sided 90% confidence interval around the mean ΔQTc estimate.* ***In case QT evaluation as described here is completed before patient studies are initiated, the level of the QTc effect and the 90% confidence interval will be used to determine the need for precautions, exclusion criteria and the level of ECG monitoring in subsequent patient trials as described in E14 2.3.*** ***Drugs with a pronounced heart rate (HR) effect*** While I agree on the point that QT evaluation conducted in patients may be informative in case the drug has a pronounced HR effect and that dose titration may be useful, it also seems important to emphasize that in most cases there will be a need for ECG monitoring in phase 3 trials based on this level of HR effect. As pointed out by the E14/S7B IWG group on several occasions, the role of the QT assessment in healthy subjects is to define which drugs would need ECG monitoring in patients, with the objective to further characterize this effect in the targeted patient population. A HR effect at this level will in most cases be known based on phase 1 studies, which further underscores the need of defining which consequences it has for subsequent patient studies. Under Decision-Making, it is stated that the sponsor can argue that a drug with this level of HR effect can be viewed as having *'low likelihood of proarrhythmic effects due to delayed repolarization',* if 6.1 requirements #1 and #3 are also met. This seems correct if the drug causes a HR increase at this level, since TdP is closely associated with bradycardia. Such drug may however, trigger coronary ischemia or worsen congestive heart failure and thereby cause life-threatening arrhythmias on this basis. A drug that causes a reduction of HR at this level, i.e., is having a strong negative chronotropic effect, is likely to cause clinically significant bradycardia and may trigger sinus pauses and AV blocks. Moreover, if the mean effect on ΔQTc is only somewhat below 10 ms, as an example, 9 ms with an UB90%CI of 17 ms, it is probably incorrect, or at least not convincingly shown, that a drug that also reduces the HR at this level should be viewed as having a *'low likelihood of proarrhythmic effects due to delayed repolarization'*. Has there been any data shared within the IWG to support this latter point? In my view, the HR example is, at best, TdP-centric, and ignores other mechanisms that clearly would warrant ECG monitoring in patients. As I see it, the example therefore leads in the wrong direction, and should be dropped from the document.

Question no.	Stakeholder no.	Comment and rationale; proposed changes
6.1	6	**Comments:** The answer to this question includes the sentence: *The design elements that include placebo and healthy participant dosing assist in decreasing variability, but their absence does not preclude interpretation.* It has been shown that the absence of placebo in a study that is analysed using concentration response analysis increases the risk of a false negative outcome (Ferber G, Sun Y, Darpo B, Garnett C, Liu J: Study Design Parameters Affecting Exposure Response Analysis of QT Data: Results From Simulation Studies. J Clin Pharmacol. 2018 May;58(5):674-685. doi: 10.1002/jcph.1065). This should be mentioned in the above sentence and interpretation needs to take this into account.
6.1	6	**Comments:** **Decision making** Under #2 of this chapter, we have *The high-quality ECG data (see ICH E14 and E14 Q&A 1) collected in the alternative QT clinical assessment do not suggest QT prolongation, generally defined as ΔQTc greater than 10 ms, as computed by the concentration-response analysis (see E14 Q&A 5.1 for details) or the intersection-union test.* The use of the term ΔQTc is misleading and should be avoided, since it suggests no correction for any placebo (diurnal) effects. This chapter introduces a new criterion to be used for labelling – in contrast to the one introduced in 2.2.4 of ICH E14 which is to be used for the planning of Phase III studies. It should be emphasised that this is an additional, distinct criterion. Moreover, the notion "greater than 10 ms" leaves open if this is a point estimate or a confidence interval, as was commonly understood up to now.
6.1	9	**Comments:** Can a recommendation for the species to be evaluated in nonclinical in vivo studies be included?
6.1	9	**Comments:** If the hERG safety margin in screening assay is very large, does the agency still believe that a repetition of assay with positive controls (a series of drugs) would be needed for consideration/ establishment of adequate exposure margin from a preclinical perspective to lower the clinical exposure margin criteria? We are concerned that this may significantly increase the burden on the sponsor in cases where it may not be necessary.

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		Proposed change: Please clarify that the establishment of exposure margin is not necessary if hERG safety margins are very large.
6.1	9	**Comments:** We would like some clarifications on the proposal for the oncology studies compared to the previous approach? Do we have to understand that the upper bound of CI for the mean change QTc <20 ms is not sufficient anymore for oncology?
6.1	9	**Comments:** "for drugs with confounding heart rate effects (i.e., >20 bpm)" This is not clear reading only available information in this Q&A: What exactly should be above 20 bpm? Does it "e.g." mean the mean change from baseline HR at any time point? **Proposed change:** We would appreciate if the EWG could provide a clearer definition for the confounding heart rate effects, if possible.
6.1	9	**Comments:** "Advanced methodologies for controlling or correcting for heart rate changes in the nonclinical in vivo studies and/or conducting QTc assessments in patients with the disease might be informative in this situation." Does it mean, that we should use either different heart rate correction in in vivo or should we be using the results from the QTc assessment of patients having the disease? In former Q&A6.1, the results from patients having the disease was the possible solution to overcome the issue of not being able to conduct essays on healthy volunteers. We understand that now it is only proposed to overcome the issue of drug induced Heart rate effect. In addition, using e.g. individual correction methods for heart rates changes in the clinical study was also proposed to overcome HR effect.
6.1	9	**Comments:** Is it possible to to detail when hERG / in vivo studies would not be performed, or are not considered a prerequisite? e.g. ICH S9? **Rationale for proposed change:** Alignment to other ICH guidelines **Proposed change:**

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		If nonclinical studies do not show low risk (or are not performed i.e. when following ICHS9)
6.1	9	**Comments:** Could you please clarify how the AEs should be presented in the database: Per study? as there will be studies with different durations, different doses/exposures, different patient populations (inclusion/exclusion criteria and as such the cardiac comorbidities will vary), monotherapy versus combinations of drugs? Healthy volunteers (if applicable) and patients with target disease separately? Number of patients/HV per Type of AE relative to total number of patients exposed? All CV SAE/AEs or just those suggestive of arrhythmia/QT prolongation? Please clarify.
6.1	9	**Comments:** Can you provide clarity on the impact/weight of each of the 3 components? Component 1 and 2 are quantitative and if they do not suggest QT prolongation, component 3, which is more difficult to quantify adequately, due to data heterogeneity and impact of comedications with possible prolonging effect, etc, should carry less weight. In the situation where the non-clinical assessment and clinical ECG assessment are clearly negative, a numerical imbalance in the AE database would not be considered to be an equivalent to either a positive non-clinical or clinical QTc assessment, nor a proarrhythmic signal. This should be clarified in the Q&A.
6.1	9	**Comments:** Please clarify point (2) in the decision making section. What would enable a low risk conclusion if points 1 and 3 are achieved, should the lack of clinical relevant QTc prolongation be defined as delta QTc <10ms with a upper bound of the CI < 20 ms
6.1	9	**Comments:** Advanced methodologies for controlling or correcting for heart rate changes in Is the nonclinical NI criterion of 10 ms used for human TQT studies applicable directly to preclinical in vivo studies and/or conducting QTc assessments in patients with the disease might? Should there be informative in this situation a different margin for nonclinical studies? Please provide more guidance/references on which methodologies, as this is not readily available on the methods to be used when HR is also affected by drug

Question no.	Stakeholder no.	Comment and rationale; proposed changes
6.1	9	**Comments:** The current paradigm seems to suffer fewer limitations when the drug in question can interact directly with ion channels proportional to drug concentration and without delay. For drugs that do not interact with ion channels the QT-prolongation assessment appears less relevant and a discussion of different drug classes e.g. highly selective drugs like peptides that may not require a TQT study would be welcome. For these compounds collecting high quality ECG at time points that may not be well defined yet in early human studies is an unnecessary burden. Non-clinically the double negative approach is challenged by the lack of scientific rationale for performing hERG test of selective peptide drugs (target specificity, size and minimal presence in cytosol at inner cavity of the hERG channel). Furthermore, many peptides have long acting profiles which does not support cross-over designs in non-rodent cardiovascular *in vivo* studies resulting in parallel study designs with reduced sensitivity and the potential need to increase group size substantially using more animals including non-human primates. To align with ICH S7B Q&A #3.2, it is suggested to clarify in ICH E14 Q&A #5.1 that the nonclinical *in vivo* assay should be conducted at exposures which cover the anticipated high clinical exposure scenario. The wording "no QTc prolongation in an *in vivo* assay" is unclear and could mean any statistically significant effect related to treatment OR an effect above a given threshold in each species. This should be clarified. While the expectation to the sensitivity in the nonclinical *in vivo* assay is clarified when used to support an integrated clinical and nonclinical risk assessment as described in ICH E14 Q&A 6.1, the same is not the case for ICH E14 Q&A 5.1. Could it be elaborated in which range the minimum detectable difference should be in order to be an assay of "sufficient sensitivity". The risk here is to not meet expectations or to use an increasing number of animals including nonhuman primates to achieve a higher than expected sensitivity.

S7B Q&A

Question no.	Stakeholder no.	Comment and rationale; proposed changes
1.1	2	**Comments:** page 11, S7B, #1.1, section 1 "If applicable, best practice considerations should be followed for assessment of additional ion channel currents (S7B Q&A 2.1)".

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		Could you be more precise and help the developers please? Are you talking about Nav1.5 and Cav1.2? Could you help on the interpretation as regards different cases: if the drug inhibits both Nav1.5 and Cav1.2 (case 1) or inhibits Nav1.5 only (case 2) or Cav1.2 (case 3)n what would be the impact?
1.1	2	**Comments:** page 11, S7B, #1.1, section 2 "the in vivo study should have sufficient power to detect a QTc prolongation effect of a magnitude similar to dedicated clinical QT studies". We suggest mentioning that the integration of continuous ECG monitoring during repeat-dose toxicology studies may now allow detection of delayed QTc effects due to drug accumulation, effects on hERG channel trafficking or hERG-blocking metabolites. However, a stand-alone in vivo telemetry study may not be needed, considering the predictive value from large animals ECG to the clinic (10 to 60 ms for the sensitivity to detect a wide range of QTc interval changes).
1.1	2	**Comments:** page 11, S7B, #1.1, 2nd § "A drug with low TdP risk would be expected to have (1) a hERG safety margin higher than the safety margins computed under the same experimental protocol for a series of drugs known to cause TdP; and (2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and human-specific major metabolites that exceed clinical exposures. If these results are used to support an integrated clinical and nonclinical risk assessment". We suggest indicating that: -a low risk in the hERG assay is defined as >30-fold ratio between the hERG-IC50 value and the clinical free therapeutic drug level -a negative nonclinical QTc evaluation is defined as >10-fold ratio between the highest free plasma drug level with no QTc effect and the relevant clinical free therapeutic drug level **Proposed change:** Provide the support for safety margin calculation and the -a low risk in the hERG assay is defined as >30-fold ratio between the hERG-IC50 value and the clinical free therapeutic drug level -a negative nonclinical QTc evaluation is defined as >10-fold ratio between the highest free plasma drug level with no QTc effect and the relevant clinical free therapeutic drug level
1.1	4 and 9	**Comments:** hERG safety margin - it would be useful to know a minimum number of compounds and state that a range of pharmacological activity is required.

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		Should some quality matrix guidance be given to exposure predictions as the quality of this prediction is important at FTiH? **Rationale for proposed change:** It could be possible just to test 2-3 TdP drugs that have a very low safety margin (i.e. <10). Against this scenario, it would be easy to make almost any compound look favourable. We propose that a cross-stakeholder expert group (under the auspices of ICH or HESI) with representation from regulators and industry, publishes a white paper defining "x" (number of compounds) and other related parameters such that the same standards are applied across all domains. **Proposed change:** (1) a hERG safety margin higher than the safety margins computed under the same experimental protocol for a series of x drugs known to cause TdP with a range of pharmacological activity
1.1	9	1. Factors that would influence the interpretation of the safety margin include the ability of the drug to block other cardiac ion channels, the potential for large excursions in clinical exposure due to intrinsic or extrinsic factors and the contributions of metabolites that inhibit the hERG channel. **Comments:** Should testing the compound effects on other cardiac channels be warranted or how the ability of the drug to block other ion channels should be considered for the hERG safety margin? If yes, do the same best practice considerations apply? More precise definition of the integrated risk assessment content and use should be given, e.g. if **and how should it be applied for non-double negative pre-clinical cases**?
1.1	9	Follow-up studies (ICH S7B Section 2.3.5) could be performed to further explore the mechanisms and assess the TdP risk. **Comments:** How do the Follow up study recommended in ICH S7B Section 2.3.5 relate to the best practice described in this Q&A? There seem to be other assays also described in the guidance. And the proarrhythmic model is not mentioned there.
1.1	9	2. In the in vivo study, the effects on the QTc interval should be assessed at exposures that cover the anticipated high clinical exposure scenario **Comments:** ICH S7A states the following on expected exposures for safety pharm studies: "Doses should include and exceed the primary pharmacodynamic or therapeutic range. In the absence of an adverse effect on the safety

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		pharmacology parameters evaluated in the study, the highest tested dose should be a dose that produces moderate adverse effects in this or in other studies of similar route and duration." This highest dose producing moderate adverse effects will likely be higher than a dose covering the high clinical exposure scenario. Can we interpret this Q&A document as such that reaching "adverse exposures" are not needed, also not for Entry into human? **Proposed change:** 2. In the in vivo study, the effects on the QTc interval should be assessed at exposures that cover the anticipated high clinical exposure scenario, **but do not need to reach moderate adverse effects**.
1.1	9	"If these results are used to support an integrated clinical and nonclinical risk assessment strategy as described in ICH E14 Q&As 5.1 & 6.1, no additional nonclinical studies are needed, except when there are factors that can confound or limit the interpretation of the nonclinical studies, such as metabolites and heart rate changes." **Comments:** The last sentence in this paragraph excludes the utility of nonclinical in vivo data in case of heart rate changes in the non-clinical studies. During the webinar on 15/16 October, Dr. Tsanugao repeatedly stated the use of the individual QT correction in nonclinical studies to counterbalance HR changes. Also, page 8 (6.1) reads: "Advanced methodologies for controlling or correcting for heart rate changes in the nonclinical in vivo studies and/or conducting QTc assessments in patients with the disease might be informative in this situation". This suggests the utility of correction for HR changes in nonclinical studies designed according to best practices. **Proposed change:** "If these results are used to support an integrated clinical and nonclinical risk assessment strategy as described in ICH E14 Q&As 5.1 & 6.1, no additional nonclinical studies are needed, except when there are factors that can confound or limit the interpretation of the nonclinical studies, such as metabolites and heart rate changes**, unless an individual QT correction is employed.**"
1.1	9	Question: What is the general strategy for use of nonclinical information as part of an integrated risk assessment for delayed ventricular repolarization and torsade de pointes that can inform the design of clinical investigations and interpretation of their results? **Proposed change:**

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		This question needs to be deleted. It is like we are trying to answer the entire updated guidance. All of this should be covered elsewhere. See additional comments. "A drug with low TdP risk would be expected to have (1) a safety margin higher than the safety margins computed under the same experimental protocol for a series of drugs known to cause TdP" How much higher? 2-fold? 10-fold? This is not helpful. "and (2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and human specific major metabolites that exceed clinical exposures." By how much? 20%? %-fold? These kinds of answers raise even more questions. We need to provide specific recommendations on safety margins like other ICH guidances (e.g., S1, S5, M3, etc,). "If the hERG assay and/or the *in vivo* QT study suggest an effect at clinical exposures, the drug has a risk of interfering with ventricular repolarization. Under…." Now we're stating the obvious and I would delete the rest of this section. It is nothing more than restating the objectives and rationale for doing these studies and all of this is covered in other sections.
1.1	9	**Comments:** The definition of "exposures…that exceed clinical exposures" should be clarified. **Rationale for proposed change:** Unless there is a clear definition of by how much the exposures in the non-clinical in vivo QT assay should exceed the clinical exposure, regulators from different regions may interpret this differently resulting in sponsors being able to use non-clinical QTc data to support a clinical concentration-response analysis in one region but not another.
1.1	9	"…(2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and human-specific major metabolites that exceed clinical exposures." **Comments:** This statement might lead to "false positive" in vivo data, e.g. in cases where a QTc prolongation was seen at 30x multiples to the human therapeutic exposure, but not any more at 3x or 10x multiples. Following the wording in this section, a QT prolongation at ANY multiples of clinical exposure would render an in vivo study positive. The same statement is true for page 9, Q 6.1. 1. and page 11, Q1.1.

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		Suggested new text: "…(2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and human-specific major metabolites that exceed clinical exposures **within a relevant range**." *OR* "….(2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and human-specific major metabolites that **cover the anticipated high clinical exposure scenario**.
1.2	1 and 9	**Comments:** For clarity it is suggested to give examples of a series of drugs with known clinical TdP risk which covers diverse electrophysiological properties or alternative clarify which electrophysiological properties should be covered.
1.2	3	**Comments:** the possibility to measure hERG inhibition in the presence of human serum is missing. In such an experimental setting actual hERG inhibition can be determined even when PPB is >99%, rather than calculating it from potentially faulty PPB determinations obtained in an independent assay.
1.2	3	**Comments:** Since there is a request to give hERG safety margins and refer to a series of drugs that have known TdP risk etc., the evaluation of the reference drugs have to be given. However, the details how these safety margins should be compared between drugs sound vague and it remains unclear how these data should be compared and presented: "….Appropriate statistical methods should be applied to quantify experimental variability and calculate uncertainty of safety margin as confidence/credible intervals…"
1.2	4 and 9	**Comments:** Methods have been developed to allow accurate determination of unbound (free) fraction <1%; in addition, the free drug hypothesis states that free plasma and free tissue concentrations are in equilibrium. **Rationale for proposed change:** Alternative methodologies have been developed to facilitate the accurate determination of plasma protein binding when conventional approaches suggest a unbound (free) **fraction** of <1%. A cross company initiative (Li et al. doi.org/10.1016/j.xphs.2017.09.005) 'indicates that values of ≤ 0.01 may be determined accurately across laboratories when appropriate methods are used.' Based on these findings we suggest that if the accuracy of methods used to determine the unbound (free) fraction when

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		experimentally determined to <1% are provided, this value can be used in calculation of the hERG safety margin. **Proposed change:** The free drug exposure is computed based on the drug's total plasma concentration and the fraction of protein binding. If this is determined to be <1%, data to support the accuracy of methods used to determine this value need to be provided. In situations where this is not possible, the unbound (free) fraction in plasma should be set to 1%.
1.2	4 and 9	**Comments:** Statement on tissues levels exceeding free plasma concentrations seems to not be in line with the free drug hypothesis. In addition, what studies should this data be based on (i.e. distribution data, QWBA, others)? **Rationale for proposed change:** A significant proportion of drugs with a high volume of distribution will have high tissue concentrations. However, the free drug hypothesis states that free plasma and free tissue concentrations should be in equilibrium; hence calculating a safety margin to the free plasma Cmax is appropriate. **Proposed change:** Margins should be calculated to free drug concentrations.
1.2	7	**Comments:** Regarding New S7B Q&As #1.2, Interpretation of hERG safety margin The hERG safety margin should be *compared to the range of safety margins computed under the same experimental protocol for a series of drugs that have known clinical TdP risk and cover diverse electrophysiological properties.* The interpretation of the safety margin of a new drug might vary due to chosen reference compounds, resulting in a high uncertainty of acceptance / interpretation by the authorities. **Proposed change:** A list of minimally requested or recommended reference compounds (including free drug exposure) should be provided. The sponsors might add additional reference compounds to improve their interpretation of the safety margin of the new drug.
1.2	9	If protein binding values cannot be accurately assessed (e.g., questionable validation of the bioanalytical method, deviations from best practices, and/or concentration-dependency of binding characteristics) or if tissue levels are likely to exceed free plasma concentrations, safety margins should be calculated for both steady-state free and total Cmax. **Comments:**

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		Does this assume providing a range and how it should be interpreted? **Proposed change:** If protein binding values cannot be accurately assessed (e.g., questionable validation of the bioanalytical method, deviations from best practices, and/or concentration-dependency of binding characteristics) or if tissue levels are likely to exceed free plasma concentrations, safety margins should be calculated **based on the appropriate/expected range**.
1.2	9	What is the recommended method to compute the hERG safety margin? **Comments:** The suggestion that submitted hERG reports must include IC50 values of known torsadogenic compounds to estimate safety margins necessitates that PK values needed for the calculation must be taken from the literature for these reference compounds. Cmax and protein binding values for drugs from literature sources vary dramatically contributing to errors in these calculations. This analysis can also be "flawed" by the choice of reference compounds and the number of IC50s run (to reduce the 90% confidence limits). All of this is especially worrisome since this safety margin computation can trigger significant added clinical work. Because of these inherent errors/issues, the requirement for this safety margin calculation could be reconsidered. We think that a weight of evidence approach taking into consideration the HERG IC50/established human free plasma levels, large animal ECG results, and Phase 1 PK/ECG data is more appropriate.
1.2	9	To assess whether the hERG block poses a risk of delaying ventricular repolarization or TdP, the resulting safety margin should be compared to the range of safety margins computed under the same experimental protocol for a series of drugs that have known clinical TdP risk and cover diverse electrophysiological properties. When a facility intends to use the model to produce data for regulatory submission, a set of control compounds should be tested to assess the consistency between the new data and the historical lab-specific validation data. **Comments**: The harmonization across regulatory agencies is needed on the recommendations/selection criteria for the list of reference drugs, the number of repeats (variability assessment) and control compounds. What are the requirements for the control compounds? Should all the compounds be re-tested with every submission or would be e.g. one compound per channel type (for hERG, Cav1.2 & Nav1.5) be enough?

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		Suggested new text: To assess whether the hERG block poses a risk of delaying ventricular repolarization or TdP, the resulting **compound effects** should be compared to the range of **effects** computed under the same experimental protocol for a series of drugs that have known clinical TdP risk in **a given compound category.** When a facility intends to use the model to produce data for regulatory submission, a **pre-defined** set of control compounds should be tested to assess the consistency between the new data and the historical lab-specific validation data.
2.1	3	**Comments:** The disadvantage of an "action potential-like" pulse protocol is that the elicited peak current is much lower than in optimized pulse protocols, leading to a diminished assay window. That's why this method is only realistic in systems with high expression level and often not applicable in automated systems. The need for a balance between "action potential-like" pulse and reasonable assay window should be mentioned.
2.1	3	**Comments:** "…After application of the test drug and if recording quality remains acceptable, a saturating concentration of a selective blocker should be applied to cells to determine residual background current…." This gives the impression that the selective blocker should be applied to test item treated cells. However, one can also block the vehicle treated cells. Is that also an acceptable approach?
2.1	3	**Comments:** Why is it recommended to use 2 or more concentrations of a positive control. If the block is sufficient with 20-80% inhibition, one concentration should be enough?
2.1	4 and 9	**Comments:** Inconstancy between 2.1. and 2.4 (line 18), 2.4 mentions the need to ensure adequate equilibration time with cells. Should the same apply to 2.1?. Please clarify inconstancy between 2.4 and 2.1. **Proposed change:** Add statement on need to ensure adequate equilibration time with cells.
2.1	7	**Comments:** Regarding New S7B Q&As #2.1, 2. Voltage protocol

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		It is asked to repeat the voltage protocols at *physiologic intervals to ensure examination and capture of frequency dependent* *effects of the test drug.* Experiments at high stimulation frequencies are much more demanding; significantly decrease success rates, or are sometimes even impossible. For example, for L-type calcium channels (CaV1.2) a frequency of 0.05 Hz is suggested in the CiPA protocol, and for hERG, a stimulation frequency of 0.2 Hz is suggested in the same protocol. However, experiments can also be performed at a stimulation frequency of 1 Hz, but this results in a lower success rate. **Proposed change:** Guidance should be given as to which frequencies will be accepted by the authorities. Since it is the aim to harmonize experimental conditions across all testing sites, at least for the hERG channel a minimal stimulation frequency should be given (similar to the assay temperature).
2.1	9	**Comments:** Comment regarding the range of safety margins computed for a series of drugs with known clinical TdP risk: - which and how many reference compounds should be taken? - should the selection of drugs be based on the indication (e.g. lower safety margins acceptable for life threatening diseases)? - which corresponding free effective Cmax values should be taken (e.g. from which literature); needs to be harmonized as well for comparison reasons **Proposed change:** - propose list of drugs and/or subset of drugs per indication together with respective free Cmax values
2.1	9	**Comments:** Comment regarding voltage protocol (it is stated "repeat at physiologic intervals"): - is the recommended voltage protocol the one that has been described in the document "recommended voltage protocols to study drug-cardiac ion channel dysfunction using recombinant cell lines – 09.18.2019"? - if yes, pulse interval would be only 5 sec which is not in concordance with the Q&A statement regarding a physiologic interval - if no, what is the recommended pulse protocol? - does this section only apply to hERG patch clamp or also for other ion channels (especially Nav1.5, Cav1.2)? **Proposed change:** - propose recommended pulse protocol in detail (same for Cav1.2 and Nav1.5)

Question no.	Stakeholder no.	Comment and rationale; proposed changes
2.1	9	"Several experimental factors are known to influence the potency of drug effects on cardiac ionic currents. " **Comments:** Would it be possible to provide a list of publications for potency differences attributed to recording temperature, composition of solutions, and the assay systems? Some of these (e.g. recording temperature) are better known than others.
2.1	9	"The voltage protocols used to evoke ionic currents should approximate the appropriate elements of a ventricular action potential and be repeated at physiologic intervals to ensure examination and capture of frequency-dependent effects of the test drug." **Comments:** Please clarify "physiologic intervals" since it is known that, for many of the currents (hERG, IKs), it may be difficult to maintain a stable baseline at frequencies that could be considered physiologic (i.e. 1 Hz). Is 0.2 Hz considered physiologic in this context?
2.1	9	"Positive and negative controls: The effects of a positive control at two or more concentrations spanning 20–80% block should be used to demonstrate assay sensitivity. " **Comments:** What are expected values for positive control effects? What is an acceptable variability range?
2.1	9	What are some "best practice" considerations when evaluating drug potency on affecting cardiac ionic currents using patch clamp method and overexpression cell lines? **Comments:** The request that every ion channel study contain plots of current amplitude, holding current, and input resistance for every cell studied is, in our opinion, unnecessary. Such requests are time consuming. Currently, GLP hERG studies (for example from CRL/ChanTest) include an example of the current amplitude versus time, an example of raw data traces, and drug inhibition at every concentration in every cell. These data should be enough to judge the quality of the submitted data. A similar comment is also valid for the need of running a full concentration response curve of a reference compound for every experiment. In our opinion, historical data from the lab, updated periodically (e.g. every 6 months) should be enough to guaranty patient's safety.

Question no.	Stakeholder no.	Comment and rationale; proposed changes
2.2	4 and 9	**Comments:** Is it worth specifying that measurements of calcium or contractility alone are not considered a substitute for measurements of the action potential? **Rationale for proposed change:** Drugs can modulate cardiac contractility through various mechanisms. As it is possible to assess cardiac action potentials in a range of cellular/tissue systems, it seems counter intuitive to rely on a surrogate measurement technology. **Proposed change:** Changes in myocyte contractions or calcium transients could have a role in further clarifying a drug's electrophysiological effects subsequently manifest as altered contractile responses if accompanied with assessment of the cardiac action potential.
2.2	9	**Comments:** Comment regarding a potential follow-up study in human cardiomyocytes: - needs this to be conducted under GLP or is GRP sufficient?
2.3	9	**Comments:** It is essential to describe the biological preparation and technology platform that impact baseline electrophysiological characteristics and drug responses. Suggest not to provide such detailed responses for each factor. The Q&A is reading like a research manuscript.
2.4	9	**Comments:** The answers are too wordy; the following edits are suggested to simplify. **Proposed change:** ~~Concentration-dependent repolarization effects can be derived based on vehicle-corrected and/or baseline subtracted comparisons of drug vs. vehicle treated preparations.~~ For higher throughput multi-well platforms, it is preferable to conduct vehicle and test drugs studies on the same plate. ….. It is important to characterize drug exposures during *in vitro* cardiomyocyte repolarization studies. ~~For well-based studies, drug exposures could be verified using media sampled from test wells or from "satellite studies" (parallel studies using identical protocols and study conditions conducted without measuring electrophysiologic measurements). With continuous flow systems the sampling of effluent from test chambers is valuable for assessing drug exposures.~~ Exposures should be presented as total drug concentration or free drug concentrations (if plasma protein binding characteristics for the media used is known).

Question no.	Stakeholder no.	Comment and rationale; proposed changes
2.5	9	**Comments:** The deleted sections below are redundant/not necessary - too detailed. **Proposed change:** ~~At minimum, it is important to characterize sensitivity to block of the prominent~~ ~~outward repolarizing current IKr/hERG with specific blocking agents (e.g., E-4031 or~~ ~~dofetilide) over relevant concentration ranges.~~ ~~Block of the inward L-type calcium current (ICaL) and late sodium current (INaL) may~~ ~~mitigate delayed repolarization. Demonstrating sensitivity to specific ICaL (e.g.,~~ ~~nifedipine or nisoldipine) and INaL (e.g., mexiletine or lidocaine) blocking agents is~~ ~~helpful for clarifying integrated cellular electrophysiological responses of multichannel~~ ~~blocking drugs.~~
3.1	4 and 9	**Comments:** There are many factors that are taken into consideration when selecting species. Should some of these in addition to alignment to toxicity studies, be included to clarify messaging? **Proposed change:** It is preferable to use the same animal species in the safety pharmacology and non-rodent toxicity studies. If an alternative species is used, this should be justified in the integrated nonclinical-clinical QTc risk assessment.
3.1	8	**Comments:** Best practice considerations for general design of the (standard) in vivo QT study should take into account relevant updates to standard practices advocated by industry (eg, social-housing telemetry is now widely available and this refinement can improve study data through improved animal welfare). **Proposed change:** Additional sentence at end: 'The use of new technologies or approaches that minimise the number of animals used and refine procedures should be employed where relevant (in line with 3Rs principles).
3.2	4 and 9	**Comments:**

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		Since there is mentioning of therapeutic and PD it would be good to clarify what PD refers to in this sentence:. An assessment of exposure in the same animals used for the pharmacodynamic assessment is encouraged. Can this be clarified? **Proposed change:** An assessment of exposure in the same animals used for the pharmacodynamic (QT) assessment is encouraged.
3.2	4 and 9	**Comments:** In this answer, the required exposure in the nonclinical QT assay is described as "should cover the anticipated high clinical exposure scenario". This is cross-referenced to Q5.1 and Q6.1. Q6.1 states that the exposure should "exceed" the clinical exposure. These answers should be consistent. Please clarify.
3.2	9	**Comments:** "exceed anticipated therapeutic concentrations" - Does this cover higher concentrations due to DDI? Is this species-dependent?
3.2	9	**Proposed change:** The deleted sentence below is not necessary. ~~In certain cases, the analysis of QTc interval together with adequate pharmacokinetic~~ ~~sampling makes it possible to perform dedicated exposure-response modeling similar to~~ ~~concentration-QT analysis for clinical QT studies.~~
3.2	9	"This could be done by sampling complete pharmacokinetic profiles in the same animals on a separate day after an adequate washout or by using limited samples from the pharmacodynamic assessment day to demonstrate consistency with full pharmacokinetic profiles generated in different animals in a separate study. " **Comments:** Do exposure assessments conducted on a separate study need to be GLP?
3.2	9	**Comments:** Advanced methodologies for controlling or correcting for heart rate changes in Is the nonclinical NI criterion of 10 ms used for human TQT studies applicable directly to preclinical in vivo studies and/or conducting QTc assessments in patients with the disease might? Should there be informative in this situation a different margin for nonclinical studies?

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		Please provide more guidance/references on which methodologies, as this is not readily available on the methods to be used when HR is also affected by drug
3.2	9	**Comments:** The current paradigm seems to suffer fewer limitations when the drug in question can interact directly with ion channels proportional to drug concentration and without delay. For drugs that do not interact with ion channels the QT-prolongation assessment appears less relevant and a discussion of different drug classes e.g. highly selective drugs like peptides that may not require a TQT study would be welcome. For these compounds collecting high quality ECG at time points that may not be well defined yet in early human studies is an unnecessary burden. Non-clinically the double negative approach is challenged by the lack of scientific rationale for performing hERG test of selective peptide drugs (target specificity, size and minimal presence in cytosol at inner cavity of the hERG channel). Furthermore, many peptides have long acting profiles which does not support cross-over designs in non-rodent cardiovascular *in vivo* studies resulting in parallel study designs with reduced sensitivity and the potential need to increase group size substantially using more animals including non-human primates. To align with ICH S7B Q&A #3.2, it is suggested to clarify in ICH E14 Q&A #5.1 that the nonclinical *in vivo* assay should be conducted at exposures which cover the anticipated high clinical exposure scenario. The wording "no QTc prolongation in an *in vivo* assay" is unclear and could mean any statistically significant effect related to treatment OR an effect above a given threshold in each species. This should be clarified. While the expectation to the sensitivity in the nonclinical *in vivo* assay is clarified when used to support an integrated clinical and nonclinical risk assessment as described in ICH E14 Q&A 6.1, the same is not the case for ICH E14 Q&A 5.1. Could it be elaborated in which range the minimum detectable difference should be in order to be an assay of "sufficient sensitivity". The risk here is to not meet expectations or to use an increasing number of animals including nonhuman primates to achieve a higher than expected sensitivity.
3.2	9	**Comments:** The wording "no QTc prolongation in an *in vivo* assay" is unclear and could mean any statistically significant effect related to treatment OR an effect above a given threshold in each species. This should be clarified.

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		To align with ICH S7B Q&A #3.2, it is suggested to clarify in ICH E14 Q&A #6.1 that the nonclinical *in vivo* assay should be conducted at exposures which cover the anticipated high clinical exposure scenario.
3.2	10	**Comments:** The section 3.2 on Page 19 states Section 3.2: "….This could be done by sampling complete pharmacokinetic profiles in the same animals on a separate day after an adequate washout…." The current statement implies taking PK samples after a washout period without re-dosing. As the assessment of drug exposure should be made relative to any effect on QT data, sampling after washout would not fulfil the requirement as exposure is expected to be lower as compared to the exposure at the effective drug level. **Proposed change:** This could be done by sampling complete pharmacokinetic profiles in the same animals dosed at the same dose levels used for the pharmacodynamic assessment on a separate day after an adequate washout
3.2 and 3.3	6	**Comments:** The answers to these questions contain a number of suggestions that need in depth statistical understanding to be put into practice appropriately. In particular, testing for equivalence or the absence of an effect is proposed without clearly stating this. I am afraid that the proposed statistics will often be understood as being based on a test for differences and will lead to unnecessary rejections of dossiers. The most salient examples are: Q 3.2: *This could be done by sampling complete pharmacokinetic profiles in the same animals on a separate day after an adequate washout or by using limited samples from the pharmacodynamic assessment day to demonstrate consistency with full pharmacokinetic profiles generated in different animals in a separate study.* "Consistency" needs to be qualified. An obvious way would be to use an equivalence test. It is not clear what variable should be tested (maximum difference, mean absolute difference, other AUC-like measures of the difference), and details including equivalence margins need to be defined. Similarities and any differences to the test for bioequivalence could be used to accomplish this. Q 3.3: *Optimally, the sponsor should demonstrate the independence of QTc to RR intervals observed in the study through QTc versus RR plots accompanied by*

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		additional information (e.g., number of matched QTc-RR pairs, correlation metric, 95% confidence intervals, p-values). Mentioning p-values here suggests a statistical test. A test for independence requires a null hypothesis and an equivalence margin to be specified properly. For the clinical situation, other methods of determining an appropriate correction method have been suggested (Tornøe CW et al (2011) Creation of a knowledge management system for QT analyses. J Clin Pharmacol 51(7):1035–1042). Although it may not be straightforward to use these in animals (sample size), an adaptation of these methods should be considered.
3.3	9	**Comments:** Individual correction methods would require multitude of drug-free measurements. Is it practical in preclinical studies?
3.4	1	**Comments:** While the expectation to the sensitivity in the nonclinical *in vivo* assay is clarified when used to support an integrated clinical and nonclinical risk assessment as described in ICH E14 Q&A 6.1, the same is not the case for ICH E14 Q&A 5.1. Could it be elaborated in which range the minimum detectable difference should be in order to be an assay of "sufficient sensitivity" for ICH E14 Q&A 5.1. The risk here is to not meet expectations or to use an increasing number of animals including nonhuman primates to achieve a higher than expected sensitivity.
3.4	8	**Comments:** A 'default' situation to improve assay sensitivity might be to increase the number of animals used within the experiments. Could an additional sentence be included at the end of the answer to encourage refinements in the animal procedures and/or alternative data processing techniques, which would be consistent with the 3Rs principles? **Proposed change:** Additional sentence at end: 'Efforts to improve assay sensitivity should ideally employ refinements to animal procedures and/or alternative data processing techniques, in preference to increasing the number of animals studied (in line with 3Rs principles).
3.4	9	**Comments:** What is the positive control to be used in preclinical studies? Is it species-dependent?
3.4	9	**Comments:** Are the limits for human TQT studies applicable to all preclinical studies?

Question no.	Stakeholder no.	Comment and rationale; proposed changes
3.4	9	**Proposed change:** The deleted sentences (preambles) are not necessary. ~~The test system used for an *in vivo* QT assay should provide a robust response. Assay~~ ~~sensitivity of relevant functional endpoints should be evaluated and reported to enable data~~ ~~interpretation (in supporting initiating first-in-human studies and/or an integrated nonclinical~~ ~~and clinical integrated risk assessment to be applied under scenarios in ICH E14 Q&As 5.1 or~~ ~~6.1) and contextualization.~~
3.5	9	"Pharmacodynamic Content Summary tables and figures showing absolute mean values, mean percent change from baseline, confidence intervals, and p-values for changes from baseline and vehicle control. Pharmacokinetic Content Tabulations of summary statistics for Cmax, AUC, and Tmax for the parent drug and metabolites along with plasma concentration vs. time plots (if sufficient samples have been collected to support their calculation)." **Comments:** 1) What does "baseline" refer to in this context? It could be the 2h predose recording on the same day of dosing, or it could refer to a 24 h recording prior to study start. In typical in vivo studies, each animal is its own control and receives every dose. Commonly, effects are compared to the time matched vehicle data from the same animal. 2) Why is AUC suggested? Typically acute functional effects are Cmax-driven, but not dependent on AUC. 3) Cmax is a measured value determined at a specific time-point. This may not reflect the true Cmax which should be determined at Tmax (blood sampling at Tmax is avoided during telemetry data capture). PK modeling may be required to estimate the Cmax. **Proposed change:** "Pharmacodynamic Content

Question no.	Stakeholder no.	Comment and rationale; proposed changes
		Summary tables and figures showing absolute mean values, mean percent change from time matched vehicle, confidence intervals, and p-values for changes from baseline and vehicle control." Pharmacokinetic Content Tabulations of summary statistics for measured Cmax, AUC, and Tmax for the parent drug and metabolites along with plasma concentration vs. time plots (if sufficient samples have been collected to support their calculation). The predicted Tmax should be indicated and the Cmax indicated (measured or modeled)".
4.1	9	**Comments:** How will conflicting conclusions from different models be addressed?
4.1	9	**Proposed change:** We suggest to delete this question. An ICH guidance is not the place to provide a scientific discourse on how to conduct experiments. If we want to keep this question in we need to simplify the answer by providing some points to consider and delete the preamble paragraph. Question: The ICH S7B guideline (Section 3.1.4) states that directly assessing the proarrhythmic risk of pharmaceuticals that prolong the QT interval would be a logical undertaking and interested parties are encouraged to develop these models and test their usefulness in predicting risk in humans. What are general principles to evaluate whether a proarrhythmic risk prediction model could be used as part of an integrated risk assessment strategy?
4.2	9	**Proposed change:** Suggest deleting this question as it is all covered elsewhere. Question: Are there any additional considerations for the use of proarrhythmia risk prediction models?
4.3	4 and 9	**Comments:** Variation in model qualification procedures between regulatory agencies is likely to be a barrier to the effective development of qualified models. **Rationale:** A common standard for the qualification of proarrhythmia models should be developed, and adopted by all regulatory agencies. This would facilitate the development of qualified models, avoid the need for multiple models for different regulatory agencies and provide a common standard for patient safety.

Question no.	Stakeholder no.	Comment and rationale; proposed changes
4.3	9	To assess whether the hERG block poses a risk of delaying ventricular repolarization or TdP, the resulting safety margin should be compared to the range of safety margins computed under the same experimental protocol for a series of drugs that have known clinical TdP risk and cover diverse electrophysiological properties. When a facility intends to use the model to produce data for regulatory submission, a set of control compounds should be tested to assess the consistency between the new data and the historical lab-specific validation data. **Comments:** The harmonization across regulatory agencies is needed on the recommendations/selection criteria for the list of reference drugs, the number of repeats (variability assessment) and control compounds What are the requirements for the control compounds? Should all the compounds be re-tested with every submission or would be e.g. one compound per channel type (for hERG, Cav1.2 & Nav1.5) be enough? **Suggested new text:** To assess whether the hERG block poses a risk of delaying ventricular repolarization or TdP, the resulting **compound effects** should be compared to the range of **effects** computed under the same experimental protocol for a series of drugs that have known clinical TdP risk in **a given compound category.** When a facility intends to use the model to produce data for regulatory submission, a **pre-defined** set of control compounds should be tested to assess the consistency between the new data and the historical lab-specific validation data.

Appendix 1

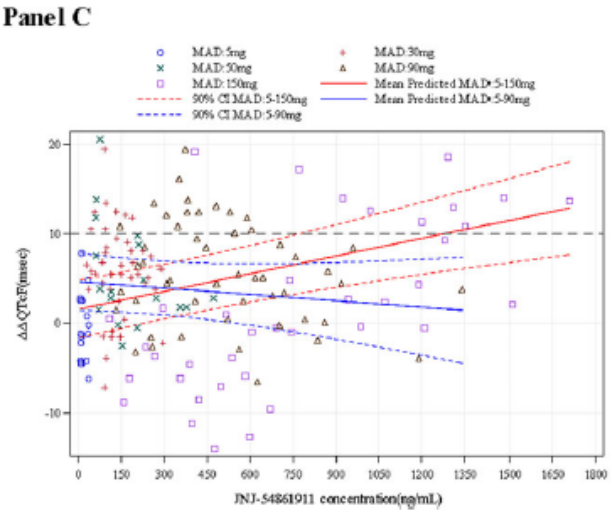

Figure 5. (A) Time course of ΔΔQTcF and JNJ-54861911 plasma concentration on day 7 across postdose timepoints after administration of 90 mg once daily to elderly subjects in the MAD study. (B) Time course of ΔΔQTcF and JNJ-54861911 plasma concentration on day 7 across postdose timepoints after administration of 150 mg once daily to young subjects in the MAD study. (C) Scatterplot of estimated ΔΔQTcF/observed plasma concentration pairs in the MAD study.
Scatterplots show all estimated ΔΔQTcF/observed plasma concentration pairs from all dose cohorts except the QT cohort. The blue lines show the predicted QT effect (ΔΔQTcF; mean ± 90%CI) using the ER model applied to cohorts 1–4 (5 to 90 mg), and the red lines show the predicted effect when adding the highest (150 mg) dose.

Source: Figure 5 in Timmers, Sinha, Darpo et al. Evaluating Potential QT Effects of JNJ-54861911, a BACE Inhibitor in Single and Multiple-Ascending Dose Studies, and a Thorough QT Trial With Additional Retrospective Confirmation, Using Concentration-QTc Analysis. J Clin Pharmacol 2018: 58; 952-64 (2).

1. Darpo B, Sarapa N, Garnett C, Benson C, Dota C, Ferber G, Jarugula V, Johannesen L, Keirns J, Krudys K, et al. The IQ-CSRC prospective clinical Phase 1 study: "Can early QT assessment using exposure response analysis replace the thorough QT study?". *Ann Noninvasive Electrocardiol.* 2014;19(1):70-81.

2. Timmers M, Sinha V, Darpo B, Smith B, Brown R, Xue H, Ferber G, Streffer J, Russu A, Tritsmans L, et al. Evaluating Potential QT Effects of JNJ-54861911, a BACE Inhibitor in Single- and Multiple-Ascending Dose Studies, and a Thorough QT Trial With Additional Retrospective Confirmation, Using Concentration-QTc Analysis. *J Clin Pharmacol.* 2018;58(7):952-64.