

ATTACHMENT 2

Responses to CHMP Topics

1. The extent to which negative nonclinical studies can exclude a clinical risk beyond reasonable doubt.

Negative nonclinical studies can exclude a clinical risk beyond a reasonable doubt. Concurrent with the issuance of Guidances S7A and S7B, it will be incumbent upon sponsors to conduct standardized *in vivo* and *in vitro* assessments that are reliable and reproducible. Thus, in the future there is an extremely low probability of introducing a QT prolonging compound into humans, which interacts with the hERG channel. This is particularly true if the margin between the doses used in nonclinical assessments are many folds higher than those used in the clinic (e.g. 100-fold+), where one would expect to see a QT effect at the higher doses. Even if nonclinical studies fail to identify the rare compound that acts independent of hERG, the compound would likely be identified in Phase 1 studies at the highest dose, if the effect on QT is significant.

Therefore, negative nonclinical and Phase 1 assessments at doses that are many folds higher than the anticipated human dose should preclude the need to conduct a thorough clinical QT study. Negative nonclinical results should trigger standard Phase 1 clinical QT assessments and selected assessments of QT prolongation in patients during Phases 2 and 3. Only if there are positive signals from nonclinical and Phase 1 studies, should a thorough clinical QT study be required.

In our experience, all development candidates that were negative in nonclinical assays did not elicit a signal in clinical studies, either QT/QTc interval prolongation or Torsade de Pointes. In over 15 years, Merck had two development candidates that prolonged QT in nonclinical assays and also increased QT/QTc intervals in humans. We recognize this is a limited data set and that one can never absolutely exclude clinical risk using non-clinical study results for any toxicological endpoint.

The value of nonclinical data for assessing risk of QT/QTc prolongation and/or risk of Torsade de Pointes is to explore mechanisms that have associated risk (e.g. hERG inhibition) and to relate the extent of delayed ventricular repolarization to the concentrations of test substance and metabolites (see Objectives of S7B Studies in ICH S7B). This information can help design the appropriate clinical strategy and assist in the interpretation of clinical data. An example comes from the briefing documents for FDA Advisory Meetings for vardenafil and alfuzosin. In both examples, there were no findings in nonclinical studies but very small QTc changes obtained in a thorough QT/QTc clinical study. Using the nonclinical data, one could conclude that QT/QTc changes were not due to a change in ventricular repolarization of the type seen with drugs known to cause Torsade de Pointes and were more likely a change in QTc secondary to vasodilatation and reflex tachycardia. In this case, ignoring the nonclinical data could lead one to over estimate the safety risk for these drug candidates.

To summarize, nonclinical data will never absolutely exclude risk but the information is very valuable in making the best overall assessment of safety. Our recommendation is that S7B nonclinical studies should be recommended and considered in both the E14 testing strategy and overall risk assessment.

2. Categories of drugs for which there would be no need for a clinical “thorough QT/QTc study.”

A thorough QT/QTc clinical study is not needed for:

- Drugs which are not systemically absorbed or which have minimal systemic exposure (e.g. ophthalmic solutions)
- Drugs in which the systemic exposure is expected to be extremely low relative to what has been safely administered to humans.
- In some situations, sponsors will have demonstrated that there is no QTc signal at clinical exposures that are significantly greater (~ > 50 fold) than the expected therapeutic exposure. This situation would typically occur in a rising single dose study for a well tolerated and safe drug prior to clear knowledge of the therapeutic exposure. In such a study, an active control would not have been administered. In this situation, if the mean difference between the drug and placebo (baseline-subtracted) for the QTc interval is < 8 ms, the utility of a thorough QT study at significantly lower exposures is questionable. In the absence of a positive control it would be possible that the study might have missed a small QT signal – however, it would be extremely unlikely that if there was a small QT signal at exposures 5 – 10 fold higher than the target therapeutic exposure, that signal would not have increased in magnitude at higher exposures and thus should have been easily detected at the highest exposure.
- For development candidates that lack a nonclinical signal, it should be sufficient to collect ECGs during Phase I studies evaluating a range of doses and from a subset of patients during the clinical program.

3. Categorization of clinical risk for drugs that prolong the mean QT/QTc interval by around 5 ms or less, 6 to 10 ms, 11 to 15 ms, 16 to 20 ms, and those that prolong the mean QT/QTc interval by more than 21 ms.

This question presumes that we can quantify a clinically and biologically meaningful human risk by grouping mean QT/QTc interval changes into 5 ms ranges. This is not the case. Rather than focusing on the interval alone to predict a meaningful clinical risk, an integrated risk assessment should be completed. It should be composed of: (1) the nonclinical signal/or lack thereof, (2) the margin between the doses used in nonclinical assessments and the anticipated human dose, and (3) the interval change in ms.

The Guidance does not state the obvious; most products do not affect the interval. Therefore, the regulatory implications of negative findings in both nonclinical and clinical evaluations are the product is deemed to have no clinically meaningful effect on repolarization.

- 4. Definition of a negative “thorough QT/QTc study” as one where the largest time-matched mean difference between the drug and placebo (baseline-subtracted) for the QTc interval is around 5 ms or less, with a one-sided 95% confidence interval that excludes an effect >8 ms. This upper bound was chosen to reflect the uncertainty related to the variability of repeated measurements.**

At Line 262, the E14 Guidance states, “Based on similar considerations, a negative ‘thorough QT/QTc study’ is one where the largest time-matched mean difference between the drug and placebo (baseline-subtracted) for the QTc interval is around 5 ms or less, with a one-sided 95% confidence interval that excludes an effect >8.0 ms.”

The Guidance should specify the research hypothesis to be supported by data from the thorough QT study (e.g., “The investigational drug does not prolong the QTc interval by more than 8 ms when compared to placebo at any of the time-points.”). How this hypothesis is tested becomes a statistical methods issue (e.g., the confidence interval excludes 8 ms). This is important since it gives the sponsor a clear assertion (hypothesis) to prove, and makes clear what Type I error rate is acceptable to regulatory agencies (e.g. 5%, as is implicit in the use of a 95% one-sided confidence interval).

Separating the statistical methodology from the hypothesis serves two purposes:

- The hypothesis to be proven is clear, and addressed first. The question comes before the answer.
- The statistical method is not pre-specified, especially without a well-understood methodology in place. For example, if the hypothesis is meant to be what is stated in the paragraph above, taking the (sample) largest time-matched mean-difference between the drug and placebo (baseline-subtracted) is *not* the optimal way to test this hypothesis.

- 5. Relative emphasis on population means value versus individual outlier analysis in determining the outcome of the “thorough QT/QTc study” as either positive or negative.**

Merck supports the draft E14 ICH guidance that emphasizes a population mean analysis as the primary analysis to determine if the outcome of the study is either positive or negative. A secondary analysis using categorical analysis of extreme values should also be performed. An "outlier analysis" (based on the formal statistical meaning, to examine a dataset for outliers that don't represent the distribution of the data) should not have any significant impact on determining the outcome of the ‘thorough QT/QTc study’. It is important to conduct an analysis of extreme values as part of the categorical analysis. All extreme values should be discussed in the context of including a comparison of frequencies (percentiles) between active treatment and placebo as well as across a range of doses. The guidance should encourage sponsors to follow all extreme values with repeat measurements and rechallenge, whenever possible. The meaning of extreme values should be determined in the context of all human data on the compound (including ECG data from studies other than the "thorough QT study") as well as the integrated nonclinical risk assessment.

6. **The extent to which results of a negative clinical “thorough QT/QTc study” can be extrapolated to exclude a risk in patients, especially in the context of patients with increased risk (e.g. extending the indication of an antihypertensive drug to include subsequently those with chronic heart failure).**

At doses that are many folds higher than the anticipated human dose, negative clinical findings in Phase 1 studies in healthy normals can be extrapolated to patients, including those at increased risk of cardiovascular disease. To predict the risk in patients, an integrated risk assessment should be composed of: (1) the nonclinical signal/or lack thereof, (2) the margin between the doses used in nonclinical assessments and the anticipated human dose, and (3) the interval change in ms.

However, when doses many fold higher than the anticipated human dose cannot be studied in normals (e.g. due to dose limiting toxicities), the extent to which negative clinical findings in healthy normals may be extrapolated to patients is problematic. In this instance, we recommend that the integrated risk assessment be compromised of: (1) the nonclinical signal/or lack thereof, (2) the interval change in ms, (3) the biological plausibility for the populations most likely to be prescribed the drug, and (4) pharmacologic class.