E14 and S7B Clinical and Nonclinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential — Questions and Answers

Guidance for Industry

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

August 2022
ICH

E14 and S7B Clinical and Nonclinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential — Questions and Answers

Guidance for Industry

Additional copies are available from:
Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov
https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs
and/or
Office of Communication, Outreach and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 71, Room 3128
Silver Spring, MD 20993-0002
Phone: 800-835-4709 or 240-402-8010
Email: ocod@fda.hhs.gov
Email: ocod@fda.hhs.gov
https://www.fda.gov/vaccines-blood-biologics/guidance-compliance-regulatory-information-biologics/biologics-guidances

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

August 2022
ICH

Contains Nonbinding Recommendations

FOREWORD

The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) has the mission of achieving greater regulatory harmonization worldwide to ensure that safe, effective, and high-quality medicines are developed, registered, and maintained in the most resource-efficient manner. By harmonizing the regulatory expectations in regions around the world, ICH guidelines have substantially reduced duplicative clinical studies, prevented unnecessary animal studies, standardized safety reporting and marketing application submissions, and contributed to many other improvements in the quality of global drug development and manufacturing and the products available to patients.

ICH is a consensus-driven process that involves technical experts from regulatory authorities and industry parties in detailed technical and science-based harmonization work that results in the development of ICH guidelines. The commitment to consistent adoption of these consensus-based guidelines by regulators around the globe is critical to realizing the benefits of safe, effective, and high-quality medicines for patients as well as for industry. As a Founding Regulatory Member of ICH, the Food and Drug Administration (FDA) plays a major role in the development of each of the ICH guidelines, which FDA then adopts and issues as guidance to industry.

Contains Nonbinding Recommendations

TABLE OF CONTENTS

Contains Nonbinding Recommendations

E14 and S7B Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential — Questions and Answers Guidance for Industry[1]

This guidance represents the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA office responsible for this guidance as listed on the title page.

INTRODUCTION

This question-and-answer (Q&A) document is intended to clarify key issues to facilitate implementing the ICH guidances for industry *E14 Clinical Evaluation of the QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs* (October 2005) and *S7B Nonclinical Evaluation of the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals* (October 2005).[2] This guidance revises ICH E14 Q&As Q12 (5.1) and Q13 (6.1), and adds new ICH S7B Q&As Q17 (1.1) to Q30 (4.2). This guidance finalizes the draft guidance issued in September 2020.

The contents of this document do not have the force and effect of law and are not meant to bind the public in any way, unless specifically incorporated into a contract. This document is intended only to provide clarity to the public regarding existing requirements under the law. FDA guidance documents, including this guidance, should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word should in Agency guidance means that something is suggested or recommended, but not required.

[1] This guidance was developed within the Expert Working Groups (Efficacy and Safety) of the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) (formerly the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) and has been subject to consultation by the regulatory parties, in accordance with the ICH process. This document has been endorsed by the ICH Assembly at *Step 4* of the ICH process, February 2022. At *Step 4* of the process, the final draft is recommended for adoption to the regulatory bodies of the ICH regions.

[2] We update guidances periodically. For the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents.

E14 QUESTIONS AND ANSWERS

I. ELECTROCARDIOGRAMS METHODOLOGY (1)[3]

Q1. Please discuss who should read electrocardiograms (ECGs), including information on the number and training of readers and the need for readers to be blinded. (1.1)

The document recommends that the reader should be skilled, but it does not identify specific training that is needed. A technician reading with a cardiologist over-reading the document would certainly be consistent with the guidance. The attempt of the guidance to limit the number of readers represents an attempt to increase consistency. The guidance asks for assessment of intra- and inter-reader variability and suggests "a few skilled readers" (not necessarily a single reader) to analyze a whole thorough QT study, since many readers may increase variability. Training would be another way to improve consistency.

It is recommended for the thorough QT Study that core ECG laboratories blind subject, time, and treatment in order to reduce potential bias. The T wave analysis, which calls for all 12 leads, can be performed after the QT analyses, and requires comparison to the baseline ECG; it can, however, be blinded as to treatment.

Q2. What is the position of ICH regarding the role of the following reading methods in the thorough QT/QTc study and other clinical trials? (1.2)

- **Fully manual**
- **Fully automated**
- **Manual adjudication (manual over-read, computer-assisted, semi-automated)?**

The techniques currently in use for the measurement of ECG intervals can be classified into three broad categories: (1) fully manual, (2) fully automated, and (3) manual adjudication. Within each of these general categories, many different methodologies are subsumed that differ in terms of lead selection, the conventions used for defining T wave offset, and the criteria for the inclusion and exclusion of U waves.

ECG readings can be performed on the following waveform presentations:

- Raw waveforms: ECG waveforms recorded from a single lead

- Representative waveforms (median beats, reference cycles): Compositional waveforms constructed by a computer-based averaging process that involves aligning and combining data from all dominant, normally conducted raw ECG waveforms from a single lead

- Global waveforms: Composite representation of cardiac electrical activity constructed by superimposing representative waveforms from all or several simultaneously recorded

[3] The numbers in parentheses reflect the organizational breakdown of the document endorsed by the ICH Assembly at Step 4 of the ICH process, February 2022.

leads to form a spatial-vector complex, by weighted averaging of individual representative complexes with low noise and long duration, or by other methods

Fully manual

When using a fully manual reading technique, a human reader is responsible for examining the ECG waveform and placing the fiducial points to mark the beginning and the end of the intervals, without the assistance of a computer algorithm. Fully manual methods of fiducial point placement can be applied to raw, representative, and global waveforms. When fully manual measurements are made from the raw ECG waveforms in a single lead, three or more cycles should be averaged, where available, to produce the final determination of interval duration. An advantage of this approach is that the reader will not be influenced by prior computer placement of the fiducial points, but a weakness can be inter- and intra-reader variability, especially when measurements are performed over an extended time period (e.g., several months). Laboratories using manual reading techniques should observe standard operating procedures based on prospectively defined criteria for determining where the fiducial points should be placed. All readers in the laboratory should be trained in the consistent application of these criteria.

Fully automated

Fully automated reading methods rely entirely upon a computer algorithm for the placement of the fiducial points and the measurement of the ECG intervals. Automated ECG interval measurements can be performed on raw, representative, or global ECG waveforms. Most digital electrocardiographs are equipped with algorithms that perform measurements on global waveforms. Although automated methods have the advantage of being consistent and reproducible, they can yield misleading results in the presence of noise or when dealing with abnormal ECG rhythms, low amplitude P or T waves, or overlapping U waves. The techniques used for construction and measurement of representative waveforms and global waveforms vary between different computerized algorithms and between different software versions within individual equipment manufacturers. As a result, between-algorithm and within-manufacturer variability of fully automated measurements can confound serial comparisons when the equipment or algorithm is not constant.

Manual adjudication (manual over-read/computer-assisted/semi-automated)

The manual adjudication approach refers to reading methods in which a computer algorithm is responsible for the initial placement of the fiducial points on the ECG waveform. A human reader subsequently reviews the algorithmic placement of the fiducial points, performing adjustments wherever the computerized measurements are considered to be inaccurate. This approach can have the advantage of greater consistency and reproducibility than fully manual readings, while providing an opportunity to correct any mistakes made by the algorithmic readings. Laboratories using manual adjudication techniques should observe standard operating procedures based on prospectively defined criteria for determining when fiducial points should be corrected. All readers in the laboratory should be trained in the consistent application of these criteria. The adjudication procedure should normally be performed on all waveforms being used

for interval determination. If an alternative approach is used, such as adjudication limited to outlier intervals above and below a reference range, this methodology should be validated as described in E14 Q&A Q3 (1.3).

The ICH E14 guidance currently recommends either fully manual or manual adjudication approaches for clinical trials in which the assessment of ECG safety is an important objective, such as the thorough QT/QTc study. When the thorough QT study is positive, fully manual or manual adjudication methods are currently recommended for an adequate sample of patients in late phase studies (see section II.C (2.3) in the E14 guidance). When the thorough QT/QTc study is negative, routine ECG safety assessments in late phase clinical trials using fully automated reading methods will be adequate.

Q3. The ICH E14 guidance contains the following statement: "If well-characterized data validating the use of fully automated technologies become available, the recommendations in the guidance for the measurement of ECG intervals could be modified." What would be expected of a sponsor that wished to validate and apply an automated reading method for regulatory submissions? (1.3)

Efforts to develop more sophisticated and reliable methods for automated ECG readings for both QT interval and T wave morphology assessment are encouraged. There are at present no large-scale studies to validate the use of fully automated reading methods in patients; however, there are examples of thorough QT/QTc studies in healthy volunteers in which automated methods have been used and validated for QT interval measurements against manual methods.

QT Interval measurement

There are at present no clear and widely accepted criteria for validation of new semi-automated or automated methods, but it is expected that each would be validated independently for its ability to detect the QT/QTc prolongation effects of drugs that are near the threshold of regulatory concern. Data supporting the validation of a new method should be submitted and could include descriptive statistics, Bland-Altman plots of agreement, superimposed plots of the baseline- and placebo-adjusted QTc and the RR as a function of time, together with data from any trials that have employed the method.

T wave morphology assessment

The suitability of automated ECG reading techniques for the assessment of morphological abnormalities has not yet been demonstrated. If a sponsor intends to develop a fully automated approach, without visual assessment for morphological changes, validation studies should include a demonstration that the automated method is capable of reading and interpreting a test set of abnormal ECGs correctly (e.g., abnormalities of T wave morphology, overlapping U waves). As with methods for QT interval determination, there are at present no clear and widely accepted criteria for validation of novel methods.

Because changes in morphology can affect interval measurement, fully manual or manual adjudication (as defined in Question Q2 (1.2)) techniques should be performed if treatment-

emergent changes in morphology are observed. If, on the other hand, no morphology changes are observed, this would support the use of automated methodologies, provided they have been validated.

Q4. How does a sponsor incorporate new technology or validate new methodology into the measurement and/or analysis of the QT interval? (1.4)

The ICH process is better suited to the determination of regulatory policy once the science in a particular area has become more or less clear. In general, it is not well-suited to the qualification or validation of new technology.

Sections II.E.1 (2.5.1) and II.E.2 (2.5.2) of the ICH E14 guidance are rather discouraging about methodology outside conventional carts and human-determined measurements. Since the ICH E14 guidance was issued, 12-lead continuous recording devices have largely supplanted cart recorders in thorough QT studies without a formal validation process because of their performance in the context of a positive control. The impact of other innovative technologies can be assessed in studies incorporating a positive control. Although some technologies could be assessed using other techniques in the absence of a positive control, this topic is more complex and beyond the scope of this question and answer.

Twelve-lead continuous recording devices and other new technologies can be used in late phase clinical trials. Even though a positive control is not used in late-stage studies, the new technology could be validated in other studies (such as the thorough QT study). In cases where a thorough QT study is not done, a sponsor can provide alternative methods for validating the technology.

Q5. The ICH E14 guidance states that QT interval corrected by Fridericia's and Bazett's correction should be submitted in all applications; is this still necessary? Is there a recommended approach to QT correction that is different from that specified in the ICH E14 guidance? (1.5)

Changes in heart rate could variably influence a drug's effect on repolarization (i.e., QT interval), and correction methods with different characteristics are often applied. The principles set below would be applicable in all clinical studies (thorough QT or other studies).

In adults, Bazett's correction has been clearly shown to be an inferior method of correcting for differences in heart rate among and within subjects. Therefore, QT interval data corrected using Bazett's corrections is no longer warranted in all applications unless there is a compelling reason for a comparison to historical Bazett's corrected QT data. Presentation of data with a Fridericia's correction is likely to be appropriate in most situations, but other methods could be more appropriate. There is no single recommended alternative (see E14 Q&A Q4 (1.4) on Incorporating New Technologies), but the following are some considerations:

1. Analyses of the same data using different models for correcting QT can generate discordant results. Therefore, it is important that the method(s) of correction, criteria for the selection of the method of correction, and rationale for the components of the method

of correction be specified prior to analysis to limit bias. Model selection should be based on objective criteria and should consider the uncertainty in parameter estimates. Alternative methods of correction should be used only if the primary method fails the prespecified criteria for selection of the method of correction.

2. Corrections that are individualized to a subject's unique heart rate QT dynamic are not likely to work well when the data are sparse or when the baseline data upon which the correction is based do not cover at least the heart rate range observed in patients on the study drug.

II. SEX (2)

Q6. There are recognized differences in the baseline QTc between men and women. These were noted in early versions of the guidance. In the E14 guidance, however, it is recommended that outliers be categorized as > 450, > 480, and > 500 milliseconds (ms), regardless of sex. Can you say why there is no sex difference in the recommendation? (2.1)

The 450, 480, and 500 ms categories refer to the values the E14 guidance suggests sponsors might use in characterizing outliers. The numbers that were specified previously for males and females referred to "normal" QTc values, which may differ for men and women. This section was not included in the final guidance, however, and such considerations would be largely irrelevant to larger durations (e.g., 480 and 500 ms). As the thorough QT/QTc study is designed to examine the propensity of a drug to prolong the QTc interval, it is appropriate to perform the study in healthy male or female volunteers.

Q7. Should we enroll both sexes in a thorough QT study, and does the study need to be powered for independent conclusions about each sex? (2.2)

Post-pubertal males have lower heart-rate corrected QT intervals than do pre-pubertal males or females generally. Women are generally smaller than men, so their exposure to a given fixed dose of a drug will generally be higher, and if a drug prolongs QT, it can be expected to prolong it more in women because of the higher exposure. It is not settled whether and how often there are sex differences in response to QT-prolonging drugs that are not explained by exposure alone.

The thorough QT study is primarily intended to act as a clinical pharmacology study in a healthy population using a conservative primary objective defining the drug's effect on QT. It is unlikely that any of a variety of baseline demographic parameters would introduce a large difference in QT response to a drug in subpopulations defined by factors such as age, co-morbidity, and sex that is not explained by exposure.

It is encouraged, but not mandatory, to include both men and women in the thorough QT study. Analyses of Concentration-Response Relationship by sex can be helpful for studying the effect of the drug on QT/QTc interval in cases where there is evidence or mechanistic theory for a sex difference. However, the primary analysis of a thorough QT study should be powered and

conducted on the pooled population. If the primary analysis is negative and if there is no other evidence suggesting sex differences, subgroup analysis by sex is not expected.

III. POSITIVE CONTROL (3)

Q8. The ICH E14 guidance emphasizes the importance of assay sensitivity and recommends the use of a positive control. In order to accept a negative thorough QT/QTc study, assay sensitivity should be established in the study by use of a positive control with a known QT-prolonging effect. Please clarify how to assess the adequacy of the positive control in the thorough QT study. (3.1)

The positive control in a study is used to test the study's ability (its *assay sensitivity*) to detect the study endpoint of interest, in this case QT prolongation by about 5 ms. If the study is able to detect such QT prolongation by the control, then a finding of no QT effect of that size for the test drug will constitute evidence that the test drug does not, in fact, prolong the QT interval by the amount of regulatory concern. There are two conditions required for ensuring such assay sensitivity:

1. The positive control should show a significant increase in QTc, i.e., the lower bound of the one-sided 95% confidence interval (CI) must be above 0 ms. This shows that the trial is capable of detecting an increase in QTc, a conclusion that is essential to concluding that a negative finding for the test drug is meaningful.

2. The study should be able to detect an effect of about 5 ms (the QTc threshold of regulatory concern) if it is present. Therefore, the size of the effect of the positive control is of particular relevance. With this aim, there are at least two approaches:

    a. To use a positive control showing an effect of greater than 5 ms (i.e., lower bound of a one-sided 95% CI > 5 ms). This approach has proven to be useful in many regulatory cases. However, if the positive control has too large an effect, the study's ability to detect a 5 ms QTc prolongation might be questioned. In this situation, the effect of the positive control could be examined at times other than the peak effect to determine whether an effect close to the threshold of regulatory concern can be detected.

    b. To use a positive control with an effect close to 5 ms (point estimate of the maximum mean difference with placebo close to 5 ms, with a one-sided 95% CI lower bound > 0). In using positive controls with smaller effects, it would be very important to have a reasonably precise estimate of the drug's usual effect.

Importantly, whatever approach is used, the effect of the positive control (magnitude of peak and time course) should be reasonably similar to its usual effect. Data suggesting an underestimation of QTc might question the assay sensitivity, thus jeopardizing the interpretability of the thorough QT study results.

Q9. Please clarify the need for blinding the positive control in the thorough QT study. (3.2)

The use of a double-blinded positive control does not appear to be essential, provided that the reading of ECGs is performed in a blinded manner as described in E14 Q&A Q2 (1.2), and the study is carefully designed to ensure that specified study procedures are followed uniformly. This means that the same protocol for administering the test drug and placebo, taking blood samples, and collecting the ECG data should also be used when giving the positive control. This does not mean that other aspects of the study, such as the duration of treatment with the positive control and the other treatment groups, would be identical. If blinding of the positive control is performed, common methods include the use of double-dummy techniques and over-encapsulation.

IV. STUDY DESIGN (4)

Q10. In the ICH E14 guidance, the recommended metric to analyze for a crossover study is the largest time-matched mean difference between the drug and placebo (baseline-adjusted) over the collection period. Please discuss the most appropriate metric to assess a drug's effect on QT/QTc interval when the data are collected in a placebo-controlled parallel design study (i.e., when there is no corresponding placebo value for each patient). (4.1)

Regardless of the study design, "the largest time-matched mean difference between drug and placebo (baseline-adjusted)" is determined as follows: The mean QTc for the drug (i.e., averaged across the study population) is compared to the mean QTc for placebo (averaged across the study population) at each time point. The "largest time-matched mean difference between drug and placebo" is the largest of these differences at any time point.

The term "baseline-adjusted" in the ICH E14 guidance implies that the baseline data are taken into account in the statistical analysis.

Differences in baseline assessment between crossover and parallel design studies are discussed in E14 Q&A Q11 (4.2).

Q11. Please discuss the need for baseline measurements, and when needed, how they should be collected, for crossover and parallel design thorough QT studies. (4.2)

Adjustment for baseline measurements is potentially useful for several purposes, including detection of carry-over effects, reducing the influence of inter-subject differences, and accounting for diurnal effects such as those due to food. There is no single best approach for baseline adjustment, but all planned baseline computations should be prospectively defined in the clinical trial protocol. Two kinds of baseline are commonly used: *time-matched* baseline (taken at exactly the same time-points on the day prior to the beginning of treatment as on the treatment day) and *predose* baseline (taken shortly prior to dosing). The *predose* baseline is used

for adjustment for inter-subject differences but not for diurnal effects. The choice of baseline is influenced by whether the study is parallel or crossover.

For a parallel-group study, a time-matched baseline allows the detection of differences in diurnal patterns between subjects that would not be detected by a predose baseline. In a parallel study, a *time-matched* baseline day, if performed, would ideally occur on the day before the start of the study.

In contrast, in a crossover study, a time-matched baseline is usually not necessary because adjustments for subject- and study-specific diurnal variation are implicit by design in the assessment of time-matched drug-placebo differences in QT/QTc effect. The *predose* baseline, therefore, is usually adequate for crossover studies.

Obtaining replicate ECG measurements (for example, the average of the parameters from about 3 ECGs) within several minutes of each nominal time point at baseline and at subsequent times will increase the precision of the estimated changes in QT/QTc effect.

V. USE OF CONCENTRATION RESPONSE MODELING OF QTC DATA (5)

Q12. The ICH E14 guidance states (in section III (3), page 12) that analysis of the relationship between drug concentration and QT/QTc interval changes is under active investigation. Has this investigation yielded a reasonable approach to concentration-response modeling during drug development? How can assessment of the concentration-response relationship guide the interpretation of QTc data? (5.1)

Concentration-response analysis, in which all available data across all doses are used to characterize the potential for a drug to influence QTc, can serve as an alternative to the by-time-point analysis or intersection-union test as the primary basis for decisions to classify the risk of a drug. In either case, this result is an important component of the totality of evidence assessment of the risk of QT prolongation. The overall assessment of risk of QT prolongation includes nonclinical data, the time course of QT prolongation, the magnitude of QT prolongation, categorical analyses of outliers, and certain adverse events in patients that can signal potential proarrhythmic effects.

There are many different types of models for the analysis of concentration-response data, including descriptive pharmacodynamic (PD) models (e.g., linear or E_{max} models), or empirical models that link pharmacokinetic (PK) models (dose-concentration-response) with PD models. It is recognized that concentration-response analyses of the same data using models with different underlying assumptions can generate discordant results. Therefore, it is important that the modeling methods and assumptions, criteria for model selection, rationale for model components, and potential for pooling of data across studies be specified prior to analysis to limit bias. Prospective specification of model characteristics (e.g., structural model, objective criteria, goodness of fit) based on knowledge of the pharmacology is recommended whenever possible. On occasion, the QT effect is not a direct function of plasma concentration. For example, drugs

that cause QT prolongation as a result of changes in protein synthesis or trafficking or drugs with accumulation into myocardial tissues might demonstrate hysteresis. Testing for model assumptions, hysteresis (a plot of data by-time point and a hysteresis loop plot), and goodness of fit should be documented.

Concentration-response analysis can be challenging when more than one molecular entity—multiple drugs or parent plus metabolites—contributes to the QTc effect.

Dose and exposure definitions

- Therapeutic dose: dose evaluated in phase 3 trial or recommended in product labeling

- Clinical exposure: mean steady-state maximum concentration ($C_{max,ss}$) associated with the maximum therapeutic dose

- High clinical exposure: exposure ($C_{max,ss}$) achieved when the maximum therapeutic dose is administered in the presence of the intrinsic or extrinsic factor (e.g., organ impairment, drug-drug interaction, food effect) that has the largest effect on increasing $C_{max,ss}$

- Supratherapeutic dose: dose that provides exposures (mean C_{max}) exceeding the high clinical scenario

Important considerations

Concentration-response data would not necessarily come from a dedicated QT study, nor would it necessarily come from a single study. However, there are several new and important considerations that are described below.

1. Data can be acquired from first-in human studies, multiple-ascending dose studies, or other studies, provided that the concentrations that can be safely achieved are well above the exposure at the maximum therapeutic dose at steady-state and reflect high clinical exposure scenario situations such as drug-drug and drug-food interactions, organ dysfunction, and/or genetically impaired metabolism. Whenever possible, sponsors are encouraged to explore a wide dose range in early-phase studies to enable characterization of effects at high concentrations.

2. Efficient concentration-response analysis using data acquired in studies with other purposes should have as much quality control as is needed for a dedicated study. This includes robust, high-quality ECG recording and analysis sufficient to support a valid assay for ECG intervals (see the E14 guidance and E14 Q&A Q1 (1.1)).

3. If there is an intention to pool ECG interval data from multiple studies, it is important to test for heterogeneity. Pooling of studies that were not planned for this purpose can produce bias. This potential should be critically discussed in the analysis plan.

4. A separate positive control would not be necessary if either of the following conditions is met:

    a. There are data characterizing the response at a sufficient multiple (commonly 2x) of the high clinical exposure (see ICH E14 guidance, section II.B.2 (2.2.2)).

    b. If the high clinical exposure has been achieved in the clinical ECG assessment but a sufficient multiple has not been obtained (e.g., for reasons of safety or tolerability, saturating absorption), then a nonclinical integrated risk assessment can be used as supplementary evidence. The reason higher doses were not tested should be adequately justified. See ICH S7B Q&A Q17 (1.1) for details; in summary, the nonclinical studies should include: (1) a hERG assay, following best practice considerations (see ICH S7B Q&As, section II (2)), that shows low risk as defined in ICH S7B Q&As Q17 (1.1) and Q18 (1.2) and (2) no evidence of QTc prolongation in an in vivo assay conducted according to ICH S7B at exposures that cover high clinical exposures (see ICH S7B Q&As, Q17 (1.1) and section III (3); note that some recommendations only apply to decision-making under ICH E14 Q&A Q13 (6.1)).

Decision-making

Both the intersection-union test and the concentration-response analysis can estimate the maximum effect of a drug treatment on the QTc interval, but they are not used to test the same hypothesis. As mentioned above, inspection of the time course of QT prolongation is important. 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.

When using a concentration-response analysis as the primary basis for decisions to classify the risk of a drug, the upper bound of the two-sided 90% confidence interval for the QTc effect of a drug treatment as estimated by exposure-response analysis should be < 10 ms at the highest clinically relevant exposure to conclude that an expanded ECG safety evaluation during later stages of drug development is not needed (see the E14 guidance, section II.B.4 (2.2.4), and E14 Q&A Q16 (7.1)).

Other uses

In addition to serving as the basis for regulatory decision-making, concentration-response analysis has established its utility in several settings enumerated below.

Providing insight into regimens not studied directly

An understanding of the concentration-response relationship can help predict the QT effects of doses, dosing regimens, routes of administration, or formulations that were not studied directly. Interpolation within the range of concentrations studied is more reliable than extrapolation above the range.

Predicting QTc effects of intrinsic and extrinsic factors that affect pharmacokinetics

Understanding the concentration-response relationship can help predict the effects of intrinsic (e.g., cytochrome P450 isoenzyme status) or extrinsic (e.g., drug-drug PK interactions) factors, possibly affecting inclusion criteria or dosing adjustments in later-phase studies.

VI. SPECIAL CASES (6)

Q13. The ICH E14 guidance states that in certain cases, a conventional thorough QT study might not be feasible. In such cases, what other methods should be used for evaluation of QT/QTc and proarrhythmic potential? (6.1)

An integrated nonclinical and clinical QT/QTc risk assessment can be particularly valuable when a thorough QT study or concentration-QTc analysis meeting similar quality control as needed for a dedicated study as described in E14 Q&A Q12 (5.1) is not feasible. This situation can arise under scenarios where a placebo-controlled comparison is not possible; 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. The design elements that include placebo and healthy participant dosing assist in decreasing variability, but their absence does not preclude interpretation.

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 Q17 (1.1) and Q18 (1.2)

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 guidance, section II.B (2.2), and E14 Q&A Q12 (5.1))

In situations where it is not possible to evaluate the QT/QTc effects at high clinical exposure, it is particularly important that the nonclinical in vivo studies are conducted at exposures covering the high clinical exposure (see ICH E14 Q&A Q12 (5.1) for definition of *high clinical exposure*).

An integrated QT/QTc risk assessment can also be particularly valuable for drugs with confounding heart rate effects that could impact accurate determination of the QTc. Advanced methodologies for controlling (e.g., pacing) or correcting for heart rate changes in the nonclinical

in vivo studies and/or conducting QTc assessments in the intended patient population might be informative in this situation. If tolerance to the chronotropic effect develops with repeat dosing, upward titration regimens can sometimes be employed to avoid or minimize the confounding effects of drug-induced heart rate changes on the QTc assessment.

Decision-Making

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.

A drug that meets the following criteria would be considered to have a low likelihood of proarrhythmic effects due to delayed repolarization:

1. The nonclinical studies, following best practice considerations for in vitro studies (see ICH S7B Q&As, section II (2)) and in vivo studies (see ICH S7B Q&As, section III (3)), show low risk as defined in ICH S7B Q&A Q17 (1.1).

2. The high-quality ECG data (see ICH E14 guidance and E14 Q&As, section I (1)) collected in the alternative QT clinical assessment do not suggest QT prolongation, generally defined under this Q&A as an upper bound of the two-sided 90% confidence interval around the estimated maximal effect on QTc less than 10 ms, as computed by the concentration-response analysis or the intersection-union test. If applicable, there should be no notable imbalances between treatment/dose arms in the proportion of subjects exceeding outlier thresholds.

3. A cardiovascular safety database that does not suggest increased rate of adverse events that signal potential for proarrhythmic effects (ICH E14 guidance, section IV (4)).

When justified, a totality of evidence argument for a drug to have a low likelihood of proarrhythmic effects due to delayed repolarization could still be made for a drug that has an upper bound of the two-sided 90% confidence interval around the estimated maximal effect on QTc of 10 ms or more. The determination will depend on the quality and details of the clinical data (e.g., estimated QTc mean and upper bound values, slope of any concentration-QTc relationship) and nonclinical data (e.g., difference between the hERG safety margin for the investigational drug and the threshold for defining low risk).

If nonclinical studies do not show low risk (or are not performed), there is reluctance to conclude a lack of an effect in an absence of a positive control; however, if the upper bound of the two-sided 90% confidence interval around the estimated maximal effect on QTc is less than 10 ms, the treatment is unlikely to have an actual mean effect as large as 20 ms.

Q14. The ICH E14 guidance does not address the approach to QT measurement during drug development in the case of combination drug products. Is it recommended that measurement of QT prolongation be performed on drug combinations? (6.2)

In general, combinations of two or more drugs are unlikely to need a thorough QT/QTc study or intensive late-stage monitoring, if the component drugs have been demonstrated to lack relevant effects in thorough QT/QTc studies as described in the ICH E14 guidance.

If one or more of the component drugs have not been individually characterized for effects on the QT/QTc interval, they may be evaluated in combination or independently.

Q15. Are sponsors expected to conduct thorough QT studies as part of the development of large proteins and monoclonal antibodies? (6.3)

Large, targeted proteins and monoclonal antibodies have a low likelihood of direct ion channel interactions and a thorough QT/QTc study is not necessary, unless the potential for proarrhythmic risk is suggested by mechanistic considerations or data from clinical or nonclinical studies.

VII. ELECTROCARDIOGRAMS MONITORING IN LATE-STAGE CLINICAL TRIALS (7)

Q16. The ICH E14 guidance describes in section II.C (2.3) (Clinical Trial Evaluation After the "Thorough QT/QTc Study") that "adequate ECG assessment to accomplish this [monitoring] is not fully established." Is there now a reasonable approach to evaluating QTc in late-stage clinical development in the case of a finding of QT prolongation prior to late phase studies? (7.1)

Clarification of approach to evaluating QTc in late-stage clinical development

The purpose of a thorough QT study is to characterize the effect of the drug on ventricular repolarization (QT interval). It is not the purpose of the thorough QT study to assess the risk of torsade de pointes (TdP) in the target population, but rather to determine whether further data are warranted to assess risk. A finding of QT prolongation above the regulatory threshold of interest (a positive thorough QT study) might call for further electrocardiographic follow-up in late phase studies. The extent of the follow-up would be affected by the magnitude of the estimated prolongation at doses and concentrations at which this occurs. If prolongation is substantial at concentrations expected to occur in clinical studies, it is important to protect patients in later trials and to obtain further information on the frequency of marked QT prolongation. In some cases in which there is a large margin of safety between therapeutic exposures and the exposures that result in significant ECG interval changes, an intensive ECG follow-up strategy might not be warranted.

The recommended intensity of the monitoring and assessment in late-stage trials will depend on:

A. The magnitude of QTc prolongation seen in the thorough QT study or early clinical studies

B. The circumstances in which substantial QT prolongation might occur (i.e., in ordinary use or only when drug concentrations are markedly increased (e.g., by renal or hepatic impairment, concomitant medications))

C. PK properties of the drug (e.g., high inter-individual variability in plasma concentrations, metabolites)

D. Characteristics of the target patient population that would increase the proarrhythmic risk (e.g., structural heart disease)

E. The presence of adverse effects that can increase proarrhythmic risk (e.g., hypokalemia, bradycardia, heart block)

F. Other characteristics of the drug (e.g., pharmacodynamics, safety pharmacology, toxicology, drug class, hysteresis)

The following examples delineate the scope of recommended ECG investigations based on outcome of the thorough QT study or early clinical studies. These could be modified by other factors such as A through F above.

Examples of ECG monitoring in late stage:

1. The thorough QT study results in a negative finding, as defined by the E14 criteria,[4] at the therapeutic dose, but the supratherapeutic dose (relative to phase 3 dose) shows mean QTc effects between 10 and 20 ms. If there is reasonable assurance that the higher dose represents drug exposures that are unlikely to be seen in the patient population, only routine ECG monitoring is recommended in late phase trials. This approach provides reassurance for safety because patients are unlikely to experience a clinically significant QTc effect.

2. The thorough QT study results in a positive finding, as defined by the E14 criteria,[5] at the therapeutic dose, with a mean prolongation < 20 ms. For drugs with this magnitude of effect on the QTc interval, intensive monitoring of phase 3 patients is called for.

   Intensive ECG monitoring in clinical trials has two main objectives. One objective is to provide protection to patients who might have large, worrisome QT intervals > 500 ms. A second objective is identifying the frequency of marked QT increases (e.g., prolonged QT > 500 ms or increases in QTc > 60 ms).

   Given the limitations of collecting ECGs in late-stage trials, the focus of the analysis is on outliers, not on central tendency. Other than descriptive statistics, detailed statistical analysis is not expected. This monitoring is intended to be performed locally, without the involvement of a central core laboratory.

[4] A *negative study,* as defined by the ICH E14 guidance criteria, is an upper one-sided 95% CI of QTc prolongation effect < 10 ms.
[5] Ibid.

The timing of ECG collection should be based on the known properties of the drug. All patients should receive baseline, steady-state, and periodic ECGs during the trial. In addition, ECGs should be collected around T_{max} at the first dose and/or around steady state in a subgroup of patients or in dedicated studies. ECG collection at around T_{max} is not important for drugs with low fluctuations between peak and trough concentrations. If the drug shows a delayed effect in QT prolongation, then the timing of ECG collection should reflect this delay.

3. The thorough QT study results in a negative finding, as defined by the E14 criteria,[6] at the therapeutic dose, but the supratherapeutic dose shows a mean effect between 10 and 20 ms. If supratherapeutic exposure is anticipated at the clinical dose only in a well-characterized subgroup, intensive monitoring as described in Example 2 above could be carried out in this subset of the phase 3 population. In this case, there should be reasonable assurance that the higher exposure is unlikely to be seen in the general patient population. In contrast, if people in the general patient population (who cannot be readily identified in advance) will in some cases achieve this higher exposure, intensive ECG monitoring in the phase 3 population is expected, as in Example 2.

4. The therapeutic dose results in a mean QTc prolongation of > 20 ms. For drugs with large QTc prolongation effects, intensive ECG assessment would be appropriate in all patients in phase 2/3. Because of the risk of TdP, another important use of ECG monitoring in late phase trials would be to assess any risk mitigation strategies (e.g., electrolyte monitoring, dose reduction strategies). Additional ECG assessment over and above what is recommended earlier in this question and answer might also be called for (e.g., 24-hour ECG recording, telemetry, multiple trough ECGs through steady state).

The sponsor is encouraged to discuss these approaches with the relevant regulatory agency or agencies prior to initiation of the phase 3 program.

S7B QUESTIONS AND ANSWERS

I. INTEGRATED RISK ASSESSMENT (1)

Q17. 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? (1.1)

The ICH S7B guidance describes a nonclinical strategy for assessing risk of delayed ventricular repolarization and QT interval prolongation (section II.C (2.3)). A mechanistic understanding of the development of torsade de pointes (TdP) and the emergence of new types of assays have made it possible to obtain more information to assess TdP risk from nonclinical assays.

[6] Ibid.

The in vitro IKr/hERG assay and in vivo QT assay as well as optional follow-up studies, as described in the ICH S7B guidance, are conducted for hazard identification and risk assessment relevant to delayed ventricular repolarization. It is generally accepted that drugs (note that the word "drug(s)" in the S7B Q&As is used interchangeably with word "pharmaceutical(s)" in the ICH S7B guidance) that delay ventricular repolarization may have increased risk of TdP.

In addition to supporting the planning and interpretation of First-in-Human clinical studies, nonclinical investigations can also contribute to an integrated risk assessment for TdP in later stages of development when clinical data are available. The following are points to consider when using in vitro IKr/hERG data and in vivo QT data in combination with clinical QT data as part of an integrated risk assessment for situations described in ICH E14 Q&As Q12 (5.1) and Q13 (6.1).

1. To predict whether or not the hERG block poses a risk of interfering with ventricular repolarization or TdP, evaluation of the hERG safety margin based on results of a best practice assay (see S7B Q&As Q18 (1.2) and Q19 (2.1)) is recommended. ICH S7A guidance, section II.F (2.6) describes considerations for when human metabolite(s) should be assessed with in vitro systems. In these cases, the metabolite's hERG safety margin should also be evaluated.

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. The adequacy of exposure to any major human metabolites should be determined (see ICH S7A guidance, sections II.C.3.b (2.3.3.2) and II.F (2.6), and S7B Q&A Q28 (3.5)). In addition, if the assay is to be used as part of an integrated clinical and nonclinical risk assessment for situations where a conventional thorough QT study is not feasible as described in ICH E14 Q&A Q13 (6.1), the in vivo study should have sufficient sensitivity to detect a QTc prolongation effect of a magnitude similar to dedicated clinical QT studies (see ICH S7B Q&A Q27 (3.4)). This additional consideration (sensitivity similar to dedicated clinical QT studies) does not apply to decision-making prior to First-in-Human studies or under ICH E14 Q&A Q12 (5.1).

A drug with low TdP risk would be expected to have (1) a hERG safety margin higher than a threshold defined based on the safety margins computed under the same experimental protocol for a series of drugs known to cause TdP (see ICH S7B Q&A Q18 (1.2) for additional details); and (2) no QTc prolongation in an in vivo assay of sufficient sensitivity conducted at exposures of parent compound and major human metabolites that exceed clinical exposures. See ICH S7A guidance, section II.F (2.6) for considerations of human metabolite(s) in the in vitro or in vivo nonclinical assays. If these results are used to support an integrated clinical and nonclinical risk assessment strategy as described in ICH E14 Q&As Q12 (5.1) and Q13 (6.1), no additional nonclinical studies are needed. When there are factors that can confound or limit the interpretation of the nonclinical studies, such as metabolites and heart rate changes, follow-up studies as described in the ICH S7B guidance (section II.C.5 (2.3.5)) can be performed to address these specific issues.

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 this scenario, the drug's TdP risk could be affected by various other factors, such as blocking of additional repolarization currents (e.g., slow delayed rectifier potassium current [Iks]), blocking of inward currents (e.g., sodium and L-type calcium currents), effects on the trafficking of ion channel proteins from cytoplasmic sites to the surface membrane, metabolites with ion channel activities, and non-ion channel mediated QT prolongation. Follow-up studies (ICH S7B guidance, section II.C.5 (2.3.5)) could be performed to further explore the mechanisms and assess the TdP risk. If applicable, best practice considerations should be followed for assessment of additional ion channel currents (S7B Q&A Q19 (2.1)), in vitro cardiomyocyte assays (S7B Q&As Q20 (2.2) to Q23 (2.5)), or in vivo studies (S7B Q&As Q24 (3.1) to Q28 (3.5)). An appropriately qualified proarrhythmia risk prediction model (see S7B Q&As Q29 (4.1) to Q30 (4.2)) could be used according to its context of use to assess the possibility of TdP in humans. Use of in vitro and in silico models can reduce animal use in follow-up studies in accordance with the 3R (reduce/refine/replace) principles. The assessment of TdP risk using these follow-up studies, although optional, can be used together with other relevant nonclinical and clinical information to contribute to the design of subsequent clinical investigations and interpretation of their results.

Q18. What is the recommended method to compute the hERG safety margin? (1.2)

A drug's potency for hERG block, usually calculated as half-inhibitory concentration (IC50), can be normalized to the drug's estimated clinically relevant exposures in patients to calculate the safety margin. As more information is obtained during the clinical development, the estimated values of clinical exposures can be refined. When estimating hERG block potency, it is recommended to use standardized procedures and to consider the principles described in S7B Q&A Q19 (2.1).

The free drug exposure is computed based on the drug's total plasma concentration and the fraction of protein binding. Because of uncertainties in the protein binding measurements, the unbound (free) fraction in plasma should be set to 1% if experimentally determined to be < 1%. This approach has been used in the regulatory assessment for the risk of drug- drug interactions. If protein binding values cannot be accurately assessed (e.g., questionable validation of the bioanalytical method, deviations from best practices), safety margins should be calculated for both steady-state free and total C_{max}.

In terms of the exposure to use in the denominator of the safety margin calculation for supporting decision-making under ICH E14 Q&As Q12 (5.1) or Q13 (6.1), it is generally recommended that the high clinical exposure be used as defined in ICH E14 Q&A Q12 (5.1) (i.e., mean steady-state maximum concentration ($C_{max,ss}$) when the maximum therapeutic dose is administered in the presence of the intrinsic or extrinsic factor that has the largest effect on increasing $C_{max,ss}$).

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 reference drugs that have known clinical TdP risk and cover diverse electrophysiological properties. Additional pharmacological principles or modeling can be used to justify the use of a certain safety margin threshold (e.g., relationship

between hERG block and QTc prolongation from semi-mechanistic pharmacokinetic/pharmacodynamic models or systems pharmacology models; see Leishman et al. 2020,[7] for examples). This should be supported by experimental data based on the principles in these S7B Q&As (e.g., same experimental protocol applied to a series of drugs with known clinical TdP risk).

Data supporting the safety margin threshold to define a drug as having low TdP risk should be supplied in or appended to the submitted study report. If a recommended hERG margin threshold is published based on principles in these S7B Q&As, a sponsor (or a contract laboratory) seeking to use the same threshold should demonstrate that the inter-laboratory variability of IC50s from a set of calibration drugs under the same experimental protocol does not significantly decrease the sensitivity of the safety margin threshold for detecting drugs that are not low risk for TdP. Appropriate statistical methods should be applied to quantify experimental IC50 variability and calculate uncertainty of the safety margin as confidence/credible intervals.

II. BEST PRACTICE CONSIDERATIONS FOR IN VITRO STUDIES (2)

Q19. What are some "best practice" considerations when evaluating drug potency on affecting cardiac ionic currents using patch clamp method and overexpression cell lines? (2.1)

As outlined in the ICH S7B guidance, the in vitro IKr/hERG assay plays a critical role in assessing the risk for delayed repolarization and QT interval prolongation prior to first administration in humans. Nonclinical investigations can also contribute to an integrated risk assessment in later stages of development when clinical QT data are available. The following "best practice" aspects should be considered when sponsors are using IKr/hERG data to support interpretation of clinical QT data in specific scenarios as described in S7B Q&As Q17 (1.1) and Q18 (1.2) and ICH E14 Q&As Q12 (5.1) and Q13 (6.1), and when using calcium (i.e., CaV1.2) and sodium (i.e., NaV1.5) data to support a proarrhythmia assessment (ICH S7B Q&A Q17 (1.1)). It is not the intent of these S7B Q&As to make specific recommendations for a sponsor's screening activities or for all IKr/hERG assays to support first administration in humans.

Several experimental factors are known to influence the potency of drug effects on cardiac ionic currents. These include the voltage protocols used to evoke specific ionic currents; experimental conditions (such as recording temperature, composition of solutions, manual versus automated assay systems); data acceptance criteria; and data analysis methods employed. Some recommended best practices are therefore provided to enhance reproducibility of in vitro results and the translation to clinical findings. These recommendations are generalizable to voltage clamp experiments characterizing potency of drug inhibition (or potentiation) of cardiac currents.

[7] DJ Leishman, MM Abernathy, and EB Wang, 2020, Revisiting the HERG Safety Margin After 20 Years of Routine hERG Scrrening, Journal of Pharmacological and Toxicological Methods, 105:106900 (https://doi.org/10.1016/j.vascn.2020.106900).

1. Recording temperature: The effects of some drugs are temperature-sensitive, and there is currently no method to predict which molecules exhibit temperature-dependent effects or the magnitude of these effects. Thus, patch clamp experiments on cells overexpressing cardiac ion channels, including hERG, CaV1.2, and NaV1.5, should be performed at near physiological temperature (35–37 °C).

2. Voltage protocol: The voltage protocols used to evoke ionic currents should approximate the appropriate elements of a ventricular action potential and be repeated at frequencies that are sufficient to minimize the possibility of missing the effects of a test drug at physiologically relevant heart rates. For hERG, a stimulation frequency of 0.2 – 1 Hertz is recommended. For CaV1.2 and NaV1.5 currents, a stimulation frequency of 0.2 Hertz is recommended. The voltage protocol should include steps that enable monitoring of cell health and consistent electrophysiological recordings throughout the experiment (i.e., estimation of input and series resistance across time). If high seal resistance is achieved, holding current and input resistance (i.e., measures of passive membrane properties at rest) can be used as indicators of cell health and experimental stability. 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 in individual cells. If prominent, background current should be factored into potency determinations.

3. Recording quality: Seal resistance should be high enough so that the leak conductance at all voltages specified by the voltage protocol and series resistance do not compromise voltage control. The extent of series resistance compensation applied to optimize voltage control should be noted. Stability of the ionic current should be demonstrated with baseline recordings (prior to drug application) of sufficient duration to characterize drug-independent changes (such as current run-down). The time course of drug effects should be monitored until steady-state effect is obtained, and each cell can be exposed to one or more drug concentrations as long as cell health and recording quality remain stable.

4. Primary endpoint measures: The primary derived endpoints are inhibitory concentration such as the IC50 value (reported in both micromolar and nanogram/milliliter units) and Hill coefficient. If 50% current inhibition could not be achieved, a justification of the highest concentration tested should be provided together with the relation of this concentration to therapeutic free and total drug levels. Where necessary, to isolate the current-of-interest, the background current remaining after a high concentration of selective blocker application should be subtracted. If current inhibition with a selective blocker could not be achieved, leak current can be calculated and subtracted from the current traces. This approach assumes that only the current-of-interest is voltage-dependent, hence evidence and justification should be provided on why it was used.

5. Data summary: Inhibition at each drug concentration for each cell should be provided, along with the mean values of IC50 and Hill coefficient (and appropriate measures of data variability). To demonstrate recording quality, the study report should also contain time-course plots of current amplitude, input resistance, holding current for individual cells in control condition followed by drug application, and drug equilibration. If time-

dependent changes such as current run-up or run-down in baseline condition were corrected for drug inhibition estimation, the correction method applied should be described.

6. Concentration verification: The concentration of test compound to which the cells were exposed should be verified by applying a validated analytical method to the solution collected from the cell chamber. Both nominal and measured concentrations should be reported. If the nominal and measured concentrations differ significantly from each other, measured concentrations should be used to construct the concentration-response relationship to estimate IC50 and Hill coefficient.

7. Positive and negative controls: The positive control drug should be one of the "reference drugs" referred to in S7B Q&A Q18 (1.2). The positive control drug should be tested using sufficient replicates and two or more concentrations achieving 20-80% block, to demonstrate consistency and reproducibility with the reference drug data. If positive control data fall outside the range of expected values, then the study is inconclusive, and it is not recommended that the data be used to support the purposes outlined in ICH E14 Q&As Q12 (5.1) and Q13 (6.1). Vehicle (negative) controls should be included in the experiments. The vehicle should include all non-compound materials in the test article solution, such as solubilizing agents and preservatives.

Q20. What are the relevant endpoints of an informative in vitro human cardiomyocyte repolarization follow-up study? (2.2)

As outlined in the ICH S7B guidance, follow-up studies (section II.C.5 (2.3.5)) can include in vitro ventricular repolarization assays. Follow-up studies are not performed for all submissions and are often designed to address specific issues. Since implementation of the ICH S7B guidance, new technologies have become available, including assays with human-induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs). S7B Q&As Q20 (2.2) to Q23 (2.5) outline best practice considerations when in vitro cardiomyocyte assays are performed as follow-up studies.

Drug-induced changes in the intracellular or extracellular action potential waveforms recorded from hiPSC-CM preparations and acutely isolated adult human ventricular myocytes reflect the integrated effect on multiple ionic currents, exchangers, and carriers. Changes in cellular repolarization recognized as markers of ventricular proarrhythmia include delayed and abnormal repolarization (manifest as early afterdepolarizations, triggered activity or irregular beating) and should be noted.

Drug-induced effects on cardiomyocyte calcium handling (calcium transients, contractility, or calcium homeostasis) could affect electrophysiological activity. These effects may manifest as abnormal mechanical activity associated with proarrhythmic mechanisms (e.g., prolonged contractions, single or multiple premature contractions linked to triggered electrical activity such as early afterdepolarizations). The assessment of changes in calcium handling is useful for describing the presence or absence of potential proarrhythmic activity. However, such measures are inadequate to fully characterize the electrophysiological effects of a drug on repolarization.

Q21. What elements of the test system should be considered for an in vitro human cardiomyocyte repolarization assay? (2.3)

It is important to describe the biological preparation and technology platform that define baseline electrophysiological characteristics and drug responses.

- Biological preparation: The origins of cells studied, and human donor characteristics should be specified. If complex preparations containing hiPSC-CMs are used (e.g., co-cultures, organoids, engineered heart tissues), descriptions of the protocols used in creating these preparations should be provided. For primary human cardiomyocyte preparations, the tissue sources, harvesting, isolation, and enrichment procedures followed should be described. Acceptable morphological and functional inclusion criteria for the preparations as well as electrophysiologic characteristics (including baseline action potential/field potential durations, spontaneous beat frequency and variability (if applicable), resting membrane potential, upstroke characteristics, conduction patterns and/or velocity) should be clearly defined. Estimates of the proportion of preparations fulfilling criteria should be included.

- Technology platform: The methodologies used (e.g., transmembrane potential recordings (whole cell patch clamp, sharp electrode, or voltage-sensing dye approaches), extracellular recordings using field potentials, visual or impedance-based motion approaches, or calcium-sensing dyes) should be clearly described. The analysis package used for marking and interpreting waveforms should be described, with representative recordings (along with pertinent waveform markings) provided. A description of the plates or chambers used (including presence or absence of flow, substrate composition, recording electrode characteristics) should be provided.

Q22. What are important considerations when designing and implementing experimental protocols for in vitro cardiomyocyte repolarization studies? (2.4)

Protocols should be designed to address a specific question (e.g., concentration-dependent effects on repolarization). The rationale of choosing a single- or sequential-dose protocol should be provided. Bath temperature should be stable at physiologic (35–37 ºC) temperature. The sampling "window" for data collection should be clearly defined. Deviations from protocols should be clearly described, along with expected consequences.

- For spontaneously beating preparations, changes in beating rate influence repolarization independent of direct drug effects on repolarizing currents. Spontaneous beat rates in the absence and presence of drugs should be clearly indicated along with the extent of drug-induced rate changes. The choice and justification of correction formula used when assessing repolarization effects in such preparations should be provided. Due to limitations of rate correction in spontaneously beating hiPSC-CMs, interpretation of potential repolarization changes may not be possible when a drug causes a rate change.

- For paced preparations, the pacing protocol (pattern and duration) should be described and performed in the presence and absence of test compound.

- To demonstrate recording quality, the study report should contain time-course plots of primary endpoints (demonstrating drug equilibration) and other parameters that can be used to infer stability of the preparations and signal recordings.

- Concentration-dependent repolarization effects can be derived based on vehicle-corrected and/or baseline subtracted comparisons of drug versus vehicle treated preparations. For higher throughput multi-well platforms, it is preferable to conduct vehicle and test drug studies on the same plate. The number of replicates (useful for evaluating reproducibility but not inferential statistical testing) should be reported. Power calculations are helpful to establish statistical sensitivity to repolarization endpoints.

- 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 and free drug concentrations (if plasma protein binding characteristics in the media used is known).

Q23. How does one define biological sensitivity of a cardiomyocyte in vitro repolarization assay? (2.5)

The electrophysiologic sensitivity of cardiomyocyte preparations should be calibrated with established positive controls to confirm their "fit for purpose" role in defining pharmacological block of cardiac ion channel(s). This is readily accomplished by constructing concentration-response curves with recognized and specific ion current blocking agents.

- 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 multi-channel blocking drugs.

III. BEST PRACTICE CONSIDERATIONS FOR THE IN VIVO QT STUDIES (3)

Q24. What are best practice considerations for species selection and general design of the (standard) in vivo QT study? (3.1)

The most appropriate species should be selected and justified (ICH S7B guidance, section II.C (3.1.3)). It is preferable to use the same animal species in the safety pharmacology and non-rodent toxicity studies to facilitate understanding of the possible relationship between adverse cardiovascular pharmacodynamic effects and structural effects on the heart, and to obtain complementary information on systemic exposure level (toxicokinetics).

Although it is customary to use conscious freely moving telemeterized animals for the in vivo QT studies, the choice of alternative model approaches (e.g., anesthetized or paced animals) might be justified in certain circumstances to achieve adequate exposures or to overcome specific compound-related challenges (e.g., changes in heart rate, tolerability, or bioavailability limitations in conscious animals).

Q25. What should be considered for exposure assessment during the in vivo QT study? (3.2)

The ICH S7B guidance states that drug exposures should include and exceed anticipated therapeutic concentrations. If the in vivo QT data are to be used as part of an integrated risk assessment for situations described in ICH E14 Q&As Q12 (5.1) and Q13 (6.1), the exposure should cover the anticipated high clinical exposure scenario (see S7B Q&A Q17 (1.1)). An assessment of exposure in the same animals used for the pharmacodynamic assessment is encouraged. Sampling should take place at relevant timepoints and in a manner that limits interference with the pharmacodynamic effects. This could be done by sampling complete pharmacokinetic profiles in the same animals on a separate day after an adequate washout or different animals. At least one pharmacokinetic sample should be obtained during the pharmacodynamic assessment day to demonstrate consistency with the full pharmacokinetic profiles. 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. This can be helpful when the study should be powered to detect an effect similar to dedicated QT studies in humans (e.g., when using in vivo QT data as part of an integrated nonclinical and clinical risk assessment as described in ICH E14 Q&A Q13 (6.1)) because it can reduce the number of animals in accordance with the 3R (reduce/refine/replace) principles. In addition, exposure-response modeling may be helpful in other circumstances when QT prolongation is observed or anticipated based on hERG assay results.

Q26. What information is needed to support the choice of heart rate correction method in an in vivo QT assay? (3.3)

Optimally, the sponsor should demonstrate the independence of QTc to RR (RR is the time between heart beats, measured as R waves on the ECG) intervals observed in the study through QTc versus RR plots accompanied by additional information (e.g., number of matched QTc-RR

pairs, correlation metric, 95% confidence intervals, p-values). QT-RR interval relationship is also important. Justification of correction factors used for QT measures should be provided when test drugs affect heart rate. In certain cases, individual QT correction based on QT-RR relationship is a preferred method because it is more accurate and sensitive than the general methods such as Bazett, Fridericia, or Van de Water when the test drugs affect heart rate. The main reason for not using correction formulae based on historical data is the fixed-rate correction coefficients. Non-rodent species show species-specific and individual differences in their QT-RR relationships.

Q27. How should the sensitivity of the assay be evaluated? (3.4)

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 the scenarios in ICH E14 Q&A Q12 (5.1) or Q13 (6.1) and contextualization. Demonstration of assay sensitivity can be achieved by defining minimum detectable differences and testing the effects of positive controls. Statistical power calculations could also be provided from historical data from the same laboratory using the identical protocol. If historical positive control data are used to justify assay sensitivity or statistical power is calculated from historical control data, then the variance of the present data should be consistent with that seen historically.

If study results are to be used to support an integrated nonclinical and clinical risk assessment described in ICH E14 Q&A Q13 (6.1), then the study should have sensitivity to detect a QTc prolongation effect of a magnitude similar to dedicated clinical QT studies, taking into consideration inter-species differences in the normal range of values for the QTc interval. The overall sensitivity of the nonclinical assay in comparison to clinical QT studies depends on both the electrocardiographic assessment and the exposure achieved in the in vivo assay relative to high clinical exposure. This can help reduce the numbers of animals used in accordance with the 3R (reduce/refine/replace) principles. The following hypothetical example is offered for consideration with recognition that the QTc threshold and exposure multiples selected for a particular study should be justified by data obtained in the specific species tested, using recognized reference compounds under conditions consistent with the best practice recommendations set forth in these S7B Q&As.

- *Hypothetical Example*: The minimal detectable difference might be 5 milliseconds if drug exposure in the animal study only covers the high clinical exposure, but might be higher if a larger multiple of high clinical exposure is achieved (e.g., 10 milliseconds if 3X high clinical exposure is achieved, or a higher QTc threshold if an even larger multiple is achieved).

Q28. What are the recommended conventions for presenting the pharmacodynamic and pharmacokinetic results of an in vivo QT assay? (3.5)

Pharmacodynamic Content

- Summary tables and figures showing absolute mean values, mean percentage change from baseline, confidence intervals, and p-values for changes from baseline and vehicle control should be included.

- If study results are being used to support ICH E14 Q&A Q13 (6.1), then data from positive controls should be included or appended. If historical positive controls are used, then the variance of the present data should be consistent with that seen historically, which can be demonstrated by reporting minimal detectable differences with by time analysis. A statement should be provided that the data for the new drug and the historical data were collected according to the same protocol and statistical analysis plan. If deviations are present, they should be clearly justified. If concentration-QTc modeling is performed, reporting should follow similar principles as for human concentration-QTc modeling (see ICH E14 Q&A Q12 (5.1)).

Pharmacokinetic Content

- Tabulations of summary statistics for C_{max}, AUC, and T_{max} for the parent drug and metabolites along with plasma concentration versus time plots (if sufficient samples have been collected to support their calculation) should be provided.

Individual animal data should be provided.

IV. PRINCIPLES FOR PROARRHYTHMIA MODELS (4)

Q29. The ICH S7B guidance (section III.D (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? (4.1)

Different models, including in silico, in vitro, ex vivo and in vivo models, have the potential to be used as part of an integrated risk assessment strategy to evaluate the proarrhythmic risk of QT-prolonging pharmaceuticals in humans. Using in vitro and in silico models can also reduce animal use in accordance with the 3R (reduce/refine/replace) principles. Because these models have a common feature of using nonclinical experimental data as input and generating human proarrhythmia risk prediction as output, they can generally be referred to as proarrhythmia risk prediction models. The model input can vary among different models, for example, ion channel pharmacology data as input to in silico models, drug-induced changes in cellular repolarization and/or arrhythmia events as input to hiPSC-CM models, and drug- induced electrocardiographic changes as input to ex vivo/in vivo models. However, the model output (either discrete risk categories or continuous risk scores) is similar among different models. Such a feature makes it possible to develop generic principles for evaluating the predictivity of proarrhythmia risk

prediction models without specifying the type of underlying experimental data used as model input. The following general principles should be applied to all proarrhythmia risk prediction models intended to be used as part of an integrated risk assessment for regulatory purposes. Although the main focus of these principles is to evaluate a model's predictivity of TdP risk, they are general enough to guide the development of models predicting different types of proarrhythmia.

1. A defined endpoint consistent with the context of use of the model. Examples of model endpoints (e.g., TdP risk versus QT prolongation risk) can be found in Li et al. 2020.[8]

2. A fully disclosed algorithm to translate experimental measurements (model input) to proarrhythmia risk (model output), allowing independent reproduction of the model development process using the associated training and validation datasets to evaluate the model performance.

3. A defined domain of applicability/scope and limitations of the model. This includes the experimental protocols to generate model input (experimental data capturing pharmacological effects of drug), and the compounds tested should have the same arrhythmic mechanisms covered by the model.

4. A prespecified analysis plan and criteria to assess model predictivity. The analysis plan should include methods to separate the training and validation steps. In the training step, a series of reference compounds is used to adjust the model. In the validation step, another series of reference compounds is used to evaluate the performance of the prespecified model. The reference compounds used for the training and validation steps should not overlap.

5. A mechanistic interpretation of the model, which describes the relationship between the model inputs and mechanism for the arrhythmia.

6. The uncertainty in the model inputs should be captured and propagated to the model predictions. The experimental variability associated with model input should be quantified using appropriate statistical methods and then translated into probabilities of the predicted risk.

After a proarrhythmia risk prediction model has been developed, a process should be followed to evaluate whether the model development complied with these principles. Such a process would support that the model is qualified for the intended context of use as part of an integrated risk assessment for regulatory purposes. Some health authorities have procedures for the formal qualification of models that allow for a model to be used within the qualified context of use without the regulatory authority needing to reconsider and reconfirm its suitability. Model developers are encouraged to contact a regulatory agency about its specific model qualification procedures. After a model has been qualified, the use of such a model is not limited to the

[8] Z Li, GR Mirams, T Yoshinaga et al., 2020, General Principles for the Validation of Proarrhythmia Risk Prediction Models: An Extension of the CiPA *In Silico* Strategy, Clinical Pharmacology & Therapeutics, 107(1): 102-111 (https://doi.org/10.1002/cpt.1647).

specific facility that submitted the qualification package. However, if another facility intends to use the qualified model, that facility should perform laboratory-specific calibration and validation of the model using a subset of the reference compounds that were originally used to develop the model. An illustrative process of performing laboratory-specific calibration and validation is given in Han et al. 2020.[9]

Q30. How can a sponsor use a model for regulatory submission and what are the limitations? (4.2)

Sponsors can use results from a qualified proarrhythmia model as one component in the totality of evidence approach to risk assessment under the context of use for which the model was developed and qualified. When a facility intends to use the model to produce data for regulatory submission, a set of control compound(s) should be tested to assess the consistency between the new data and the historical laboratory-specific validation data. The number and type of laboratory-specific calibration and control compounds should be justified.

If a proarrhythmia model is included in a regulatory submission, proof of qualification of the model under the guidance of the general principles in this Q&A should be provided in an appendix to the study report. Supportive documentation could include published papers, if the included validation dataset is described in sufficient detail to allow an independent assessment. Importantly, the general principles for model qualification set forth in this Q&A only support the use of a proarrhythmia risk prediction model as part of an integrated risk assessment that incorporates all relevant nonclinical and clinical information.

[9] X Han, M Samieegohar, BJ Ridder et al., 2020, A General Procedure to Select Calibration Drugs for Lab-Specific Validation and Calibration of Proarrhythmia Risk Prediction Models: An Illustrative Example Using the CiPA Model, Journal of Pharmacological and Toxicological Methods, 105:106890 (https://doi.org/10.1016/j.vascn.2020.106890).