Guidance for Industry

E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs

---

Questions and Answers

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)

November 2008
ICH

Guidance for Industry

E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs

---

Questions and Answers

Additional copies are available from:

Office of Communications
Division of Drug Information, HFD-240
Center for Drug Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave.
Bldg. 51, Room 2201
Silver Spring, MD 20993-0002
(Tel) 301-796-3400
http://www.fda.gov/cder/guidance/index.htm

Office of Communication, Training and
Manufacturers Assistance, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Rockville, MD 20852-1448
http://www.fda.gov/cber/guidelines.htm.

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)

November 2008
ICH

TABLE OF CONTENTS

Contains Nonbinding Recommendations

Guidance for Industry[1]

E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs — Questions and Answers

This guidance represents the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

I. INTRODUCTION

Since the E14 guidance was finalized in 2005, experiences with implementation within the ICH regions have given rise to requests for clarification. This question and answer (Q&A) document is intended to clarify key issues.

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. QUESTIONS AND ANSWERS

Q1: The 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" (TQT study), assay sensitivity should be established in the study by use of a positive control

[1] This guidance was developed within the Expert Working Group (Efficacy) of 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 Steering Committee at *Step 4* of the ICH process, June 2008. At *Step 4* of the process, the final draft is recommended for adoption to the regulatory bodies of the European Union, Japan, and the United States.

with a known QT-prolonging effect. Please clarify how to assess the adequacy of the positive control in the TQT study.

A1: 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 milliseconds (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[2] 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 been 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 TQT study results.

[2] In this document, the terms *require*, *must*, and *need* refer to scientific necessity, not legal necessity. This Q&A guidance offers additional information for implementing the recommendations in ICH E14 and is not intended to create any new expectations beyond current regulatory requirements. The contents of the document are guidance only and do not impose any requirements on readers or on the FDA.

Q2: Please discuss who should read ECGs, including the number and training of readers and the need for readers to be blinded.

A2: The document recommends that the reader should be skilled but does not identify specific training that is needed. A technician reading with cardiologist over-read would certainly be consistent with the guidance. The attempt of the guidance to limit the number of readers represented 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.

Q3: There are recognized differences in the baseline QTc between men and women. These were noted in early versions of the guidance. In E14, however, it is recommended that outliers be categorized as >450, >480, and >500ms, regardless of gender. Can you say why there is no gender difference in the recommendation?

A3: The 450, 480, and 500 ms categories refer to the values the E14 document suggests sponsors might use in characterizing outliers. The numbers previously specified for males and females referred to "normal" QTc values, which may differ for men and women. This section was not included in the final document, however, and such considerations would be largely irrelevant to larger durations (e.g., 480, 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 male or female healthy volunteers.

Q4A: What is the position of ICH regarding the role of the following reading methods in the Thorough QT/QTc study and other clinical trials?

- ***fully manual***
- ***fully automated***
- ***manual adjudication (manual over-read, computer-assisted, semi-automated)***

A4A: The techniques currently in use for the measurement of ECG intervals can be classified into three broad categories: fully manual, fully automated, and 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 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 question 4B.

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 E14 document). When the Thorough QT/QTc study is negative, routine ECG safety assessments in late phase clinical trials using fully automated reading methods would be adequate.

Q4B: 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?

A4B: 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 4A) 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.

Q5: ***In E14, the recommended metric to analyze for a cross-over 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).***

A5: 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 E14 implies that the baseline data are taken into account in the statistical analysis.

Differences in baseline assessment between cross-over and parallel design studies are discussed in Question 6.

Q6: ***Please discuss the need for baseline measurements (and when needed, how they should be collected) for cross-over and parallel design TQT studies.***

A6: 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 "pre-dose" baseline (taken shortly prior to dosing). The "pre-dose" 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 cross-over.

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 cross-over 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 "pre-dose" baseline is therefore usually adequate for cross-over studies.

Obtaining replicate ECG measurements (for example, the average of the parameters from about three 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.

Q7: Please clarify the need for blinding the positive control in the TQT study.

A7: 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 question 4A 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.