Translational Models and Tools to Reduce Clinical Trials and Improve Regulatory Decision Making for QTc and Proarrhythmia Risk (ICH E14/S7B Updates)

David G. Strauss[1,*], Wendy W. Wu[1], Zhihua Li[1], John Koerner[2] and Christine Garnett[3]

After multiple drugs were removed from the market secondary to drug-induced torsade de pointes (TdP) risk, the International Council for Harmonisation (ICH) released guidelines in 2005 that focused on the nonclinical (S7B) and clinical (E14) assessment of surrogate biomarkers for TdP. Recently, Vargas *et al*. published a pharmaceutical-industry perspective making the case that "double-negative" nonclinical data (negative *in vitro* hERG and *in vivo* heart-rate corrected QT (QTc) assays) are associated with such low probability of clinical QTc prolongation and TdP that potentially all double-negative drugs would not need detailed clinical QTc evaluation. Subsequently, the ICH released a new E14/S7B Draft Guideline containing Questions and Answers (Q&As) that defined ways that double-negative nonclinical data could be used to reduce the number of "Thorough QT" (TQT) studies and reach a low-risk determination when a TQT or equivalent could not be performed. We review the Vargas *et al*. proposal in the context of what was contained in the ICH E14/S7B Draft Guideline and what was proposed by the ICH E14/S7B working group for a "stage 2" of updates (potential expanded roles for nonclinical data and details for assessing TdP risk of QTc-prolonging drugs). Although we do not agree with the exact probability statistics in the Vargas *et al*. paper because of limitations in the underlying datasets, we show how more modest predictive value of individual assays could still result in low probability for TdP with double-negative findings. Furthermore, we expect that the predictive value of the nonclinical assays will improve with implementation of the new ICH E14/S7B Draft Guideline.

On August 27, 2020, the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) released the Draft Guideline "E14 and S7B Clinical and Nonclinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential—Questions and Answers."[1] The Draft ICH Guideline, which contains revised Questions and Answers (Q&As) to the E14 clinical guideline and new Q&As to the S7B nonclinical guideline, was subsequently opened for public comment in regulatory regions around the world.[2] To raise public awareness of the ICH E14/S7B Draft Guideline (also referred to as the new/draft Q&As in this paper), the ICH E14/S7B working group presented a 2-day webinar on October 15–16, 2020.[3] The webinar covered the background, motivation for, and overview of the new E14/S7B Q&As, followed by presentations on each of the main Q&A topics and example cases that highlighted the potential impact of applying the principles in the new Q&As on reducing clinical trials and assisting regulatory evaluation.[3] In addition, there was an hour of live response to questions submitted by the audience from a panel of ICH E14/S7B working group members that included representatives from every regulatory region and industry group. With 2,355 registered attendees from 69 countries, there was substantial interest in the topic. The presentations and complete recordings are available for download on the U.S. Food and Drug Administration (FDA) website.[3]

Prior to release of the new ICH E14/S7B Draft Guideline, Vargas *et al*.[4] submitted an article to *Clinical Pharmacology & Therapeutics* titled "Time for a Fully Integrated Nonclinical–Clinical Risk Assessment to Streamline QT Prolongation Liability Determinations: A Pharma Industry Perspective." The article was written by 35 pharmaceutical industry authors from the "ICH S7B-E14 Industry Support Group," 8 of whom are industry representatives on the ICH E14/S7B working group. The group's objective was to lay out the state of present and emerging science that advocate for the integration of the ICH S7B (nonclinical) and E14 (clinical) strategies and practice. The authors also presented new analyses of publicly available data on the predictive value of the ICH S7B "core assays"

[1]Division of Applied Regulatory Science, Office of Clinical Pharmacology, Office of Translational Sciences, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA; [2]Division of Pharm/Tox for Cardiology, Hematology, Endocrinology and Nephrology, Office of Cardiology, Hematology, Endocrinology and Nephrology, Office of New Drugs, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA; [3]Division of Cardiology and Nephrology, Office of Cardiology, Hematology, Endocrinology and Nephrology, Office of New Drugs, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA. *Correspondence: David G. Strauss (david.strauss@fda.hhs.gov)

Review of ICH E14/S7B Updates and Vargas *et al*. "Time for a Fully Integrated Nonclinical–Clinical Risk Assessment to Streamline QT Prolongation Liability Determinations: A Pharma Industry Perspective"

Received December 9, 2020; accepted December 14, 2020. doi:10.1002/cpt.2137

(*in vitro* hERG and *in vivo* heart-rate corrected QT (QTc) assays) individually and in combination. The authors concluded that "It is our collegial opinion that if the emergent Q&As are to have a meaningful positive impact on drug development, some key points need to be addressed and considered by the [ICH E14/S7B working group]." The four key points raised were:

1. A High Prevalence of Negative "Thorough QT" (TQT) Studies
2. Conduct of the TQT Study Can be Confounded or Not Feasible
3. Multiple Ion Channel Blockade
4. Direct vs. Indirect Effects on QTc Prolongation

Vargas *et al*.[4] referred only to the ICH E14/S7B working group Concept Paper,[5] as the new ICH E14/S7B Draft Guideline[1] had not been released by the time of paper acceptance for publication. The authors of this article are the ICH E14/S7B working group members from the FDA. We review the data presented and points raised by Vargas *et al*.[4] within the context of what has been addressed in the recently released ICH E14/S7B Draft Guideline[1] ("stage 1" of Q&As as described in the ICH E14/S7B Concept Paper)[5] and what the ICH E14/S7B working group has indicated it could address in a "stage 2" of Q&As to ICH E14/S7B.[5,6]

Background, motivation, and overview of updates to ICH E14 and S7B

As implemented in 2005, ICH S7B[7] nonclinical safety pharmacology studies are conducted to inform safety before first-in-human dosing. As Vargas *et al*.[4] pointed out, "In contrast with the standardized execution and expectations for the TQT study, there are no standard protocols, experimental conditions, or regulatory expectations to guide the execution of the nonclinical core assays." As a result, nonclinical data exhibit a wide degree of data variability, stemming from diverse experimental designs and data quality control metrics specific to individual laboratories. The FDA uses these nonclinical data to identify investigational drugs with large effects on repolarization before first-in-human dosing, and the final clinical assessment generally relies on ICH E14[8] assessment of human QTc, which has the sensitivity to detect small repolarization delay of ~ 5 milliseconds (ms). However, both the ICH S7B[7] and E14[8] highlight the need for integration of information in a manner that is informative as a totality of evidence approach and an integrated risk assessment can consider the combined predictive value of multiple assays (i.e., predictive tests). In 2018, the ICH E14/S7B working group reached an agreement on a 2-stage approach to update both documents.[5] The "stage 1" of Q&A updates is currently in step 3 of the ICH process, which, after releasing a Draft ICH Guideline, involves consulting the public, discussing the comments, and revising (as needed) and finalizing the ICH Guideline.

VALUE PROPOSITION OF THE REVISED E14 Q&AS AND NEW S7B Q&AS: LEVERAGING AN INTEGRATED NONCLINICAL AND CLINICAL RISK ASSESSMENT

For ICH E14, the revised Q&As provide recommendations for how "double-negative" nonclinical data from the S7B core assays (i.e., low risk for *in vitro* hERG block and *in vivo* QTc prolongation as defined in the E14/S7B draft Q&As) can be used to inform regulatory decision making in late-stage clinical development and at the time of a marketing application.[1] **Figure 1** shows the breakdown of QT study reports submitted to the FDA from 2016–2020: 44% with conventional E14 TQT studies, 32% with (phase I) concentration-QTc data (E14 Q&A 5.1) and 24% using alternative study designs when a TQT or equivalent is not possible (E14 Q&A 6.1).[9] The potential value of the revised E14 Q&As 5.1 and 6.1 in combination with the new S7B Q&As is to streamline drug development by leveraging nonclinical data to reduce the number of clinical trials (TQT studies) and improve regulatory decision making to reach a low-risk determination at the time of a marketing application when a TQT study or equivalent cannot be performed.

Figure 2 shows the diagram from the original S7B Guideline in the center (**Figure 2a**) and where the new draft Q&As to S7B (**Figure 2b–d**) and revised draft Q&As to E14 (**Figure 2d**) fit into the integrated testing strategy. This highlights how the core assays of *in vitro* hERG and *in vivo* QTc are considered as a part of the integrated risk assessment. In general, when these assays are both negative, additional follow-up studies are not required, and this supports moving forward with first-in-human studies.[7] The new integrated risk assessment draft Q&As (S7B Q&As 1.1–1.2; **Figure 2d**) describe how nonclinical data are used prior to human testing as primarily described in the original S7B Guideline, and how they can be used later in clinical development to support regulatory decision making.[1] Specific recommendations (**Figure 2d**) are put forward for the use of double-negative nonclinical data to inform the revised E14 draft Q&As 5.1 and 6.1, and more general recommendations are put forward for how to assess non-double-negative scenarios (i.e., when hERG block or QTc prolongation is present).

To increase regulatory confidence for using nonclinical data for regulatory decision making during clinical development and at the time of a marketing application, S7B Q&As were developed to address the lack of standardization and lack of explicit statements of regulatory expectations in S7B that Vargas *et al*.[4] pointed out. These Q&As included best practice considerations for the S7B core assays (**Figure 2b**: hERG assay – S7B Q&A 2.1; *in vivo* QTc assay – S7B Q&As 3.1–3.5), and for potential follow-up studies, including additional cardiac ion channel assays (**Figure 2b**: S7B Q&A 2.1) and *in vitro* cardiomyocyte assays (**Figure 2c**: S7B Q&As 2.2–2.5).[1] In addition, principles for the use of proarrhythmia models (**Figure 2c**: S7B Q&As 4.1–4.3) were developed to provide guidance on how proarrhythmia models (including *in silico*, *in vitro*, *ex vivo*, or *in vivo* models) can be used as part of an integrated risk assessment (**Figure 2d**) strategy to evaluate the proarrhythmic risk of QTc-prolonging drugs in humans.[1] The S7B integrated risk assessment Q&As also describe how different follow-up studies and proarrhythmia models can be used if the S7B core assays are confounded (e.g., large changes in heart rate in *in vivo* studies) or to assess torsade de pointes (TdP) risk for a drug that blocks hERG and/or prolongs QTc. Furthermore, additional details are provided on FDA's website (e.g. recommended ion channel protocols)[10] and in a series of scientific white papers: a systematic strategy for

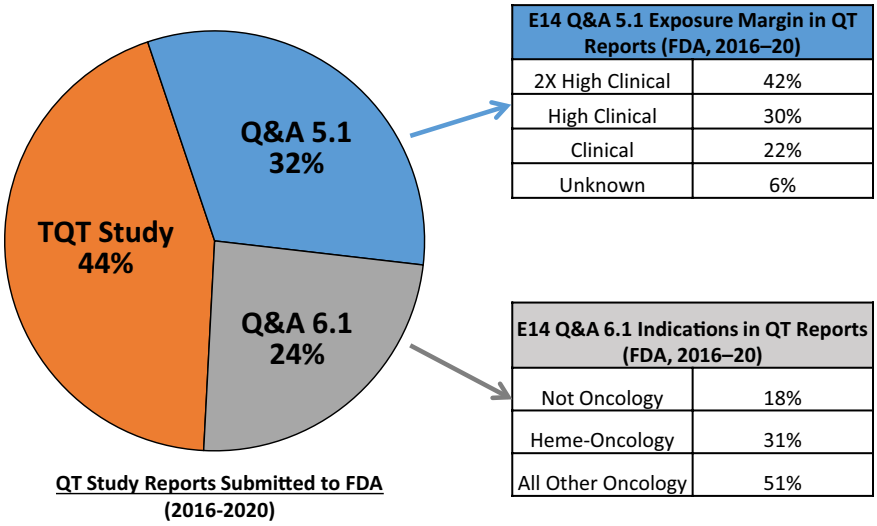


Figure 1 Types of clinical QT study report submitted to the FDA from 2016 to 2020: 44% were conventional thorough QT (TQT) studies, 32% relied on concentration-QTc analysis as described in ICH E14 Q&A 5.1 (typically from phase I clinical trials), and 24% utilized alternative study designs described in E14 Q&A 6.1 (typically without placebo, positive control, and/or reaching exposure levels required for TQT or E14 Q&A 5.1). For study reports under E14 Q&A 5.1, only 42% reached 2 times high clinical exposure level, which has been the requirement under E14 Q&A 5.1 for waiving the need for positive control in a TQT study. Under the revised draft Q&A 5.1, with double-negative nonclinical data the exposure will only need to reach high clinical exposure to waive the need for a positive control, allowing for more TQT substitutes. For study reports with alternative study designs (E14 Q&A 6.1), 82% were for oncology indications. Under the existing Q&A 6.1, studies can only reach a conclusion of "no large QTc effects." Under the revised draft Q&A 6.1, when combined with double-negative nonclinical findings, a conclusion of low likelihood of proarrhythmic effects due to delayed repolarization can be reached. FDA, US Food and Drug Administration; Q&A, question and answer.

estimating hERG block potency,[11] best practices for human stem cell-derived cardiomyocyte assays,[12] general principles for the validation of proarrhythmia risk prediction models,[13] and a general procedure for selecting drugs to calibrate a proarrhythmia model when implementing it at different laboratories.[14]

USE OF DOUBLE-NEGATIVE NONCLINICAL DATA - COMMENTS ON PROPOSAL BY VARGAS *ET AL*.

Going beyond what is contained in the new E14/S7B draft Q&As, Vargas *et al*.[4] proposed that the double-negative nonclinical findings in a "robust, comprehensive and integrated ICH S7B data package" could more broadly be used to reduce the need for detailed QT-focused clinical evaluation for small molecule drugs. To support their proposal, they present a new integrated analysis of published nonclinical-clinical datasets with the following definitions:

- The hERG assay is considered negative when the hERG safety margin is ≥ 30-fold. The hERG safety margin is defined as the hERG concentration of drug producing 50% inhibition (IC_{50}) divided by the estimated clinical free therapeutic plasma concentration.
- The *in vivo* QTc assay is considered negative when the ratio between the highest free plasma concentration without effect (e.g., < 10 milliseconds prolongation of QTc) and the estimated clinical free therapeutic exposure is ≥ 10-fold. If this ratio is < 10, the outcome is positive.

In the supplement, Vargas *et al*. average the discriminatory power (sensitivity and specificity) of these nonclinical assays across multiple literature studies for the separate endpoints of clinical QTc prolongation (hERG studies[15–19]; *in vivo* QTc studies[15,16,18,20]) and TdP (hERG studies[19,21–24]; *in vivo* QTc studies[23,25,26]). In order to use the sensitivity and specificity of these assays to predict the probability of clinical QTc prolongation or TdP, they estimate the prior (pre-test) probability of each clinical endpoint and convert sensitivity and specificity of each assay to likelihood ratios (see **Figure 3a** for examples). By doing this, the probability of the clinical endpoints can be determined after having the results of a hERG assay or *in vivo* QTc assay alone (**Figure 3b**), or when combining them (**Figure 3c,d**). Based on their new analysis, Vargas *et al*. conclude that double-negative nonclinical data result in a "very low probability of clinical QTc prolongation and TdP risk."[4] While we agree that double-negative hERG and *in vivo* QTc assays suggest low probability of clinical QTc prolongation and TdP, we caution about citing the exact probability values as stated by Vargas *et al*. Next, we discuss assumptions made by Vargas *et al*. to derive these numbers, limitations in the cited datasets, and how the predicted probabilities change with sensitivity analyses or alternative data sources.

Prior probability of clinical QTc prolongation and TdP liability

The Vargas *et al*.[4] supplemental material describes their rationale for obtaining a pre-test probability of 20% for QTc prolongation and 10% for TdP. With regard to QTc prolongation, the 20% number is consistent with a recent analysis of the FDA's database of clinical QT study reports from 2016–2020, which revealed that 19% were positive (**Figure 4**).[6] However, the FDA database does not capture drugs where a QTc signal was seen in a phase I study or TQT study where the sponsor discontinued development before

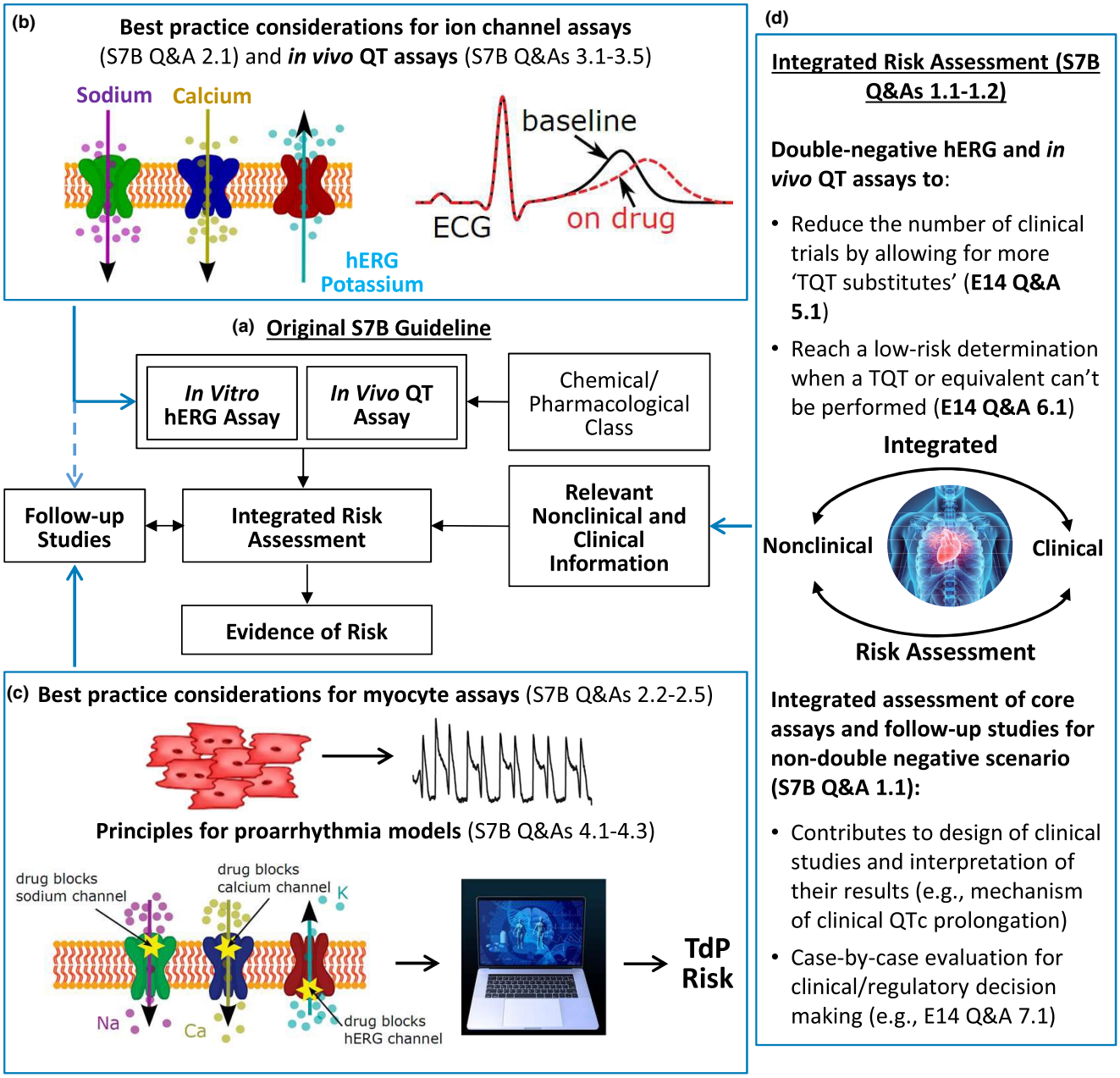


Figure 2 Schematic diagram of how the new ICH E14/S7B draft Q&As fit into the original S7B Guideline. (**a**) Diagram from the original S7B Guideline. (**b**) New S7B Q&As on best practice considerations for the core S7B assays (hERG and *in vivo* QTc) and additional ion channel assays that may be used as follow-up studies. (**c**) New S7B Q&As on best practice considerations for *in vitro* cardiomyocyte assays and principles for proarrhythmia models. (**d**) The new S7B integrated risk assessment Q&As in combination with the revised E14 Q&As describe how nonclinical data can be used to reduce the number of thorough QT (TQT) studies and reach a low-risk determination when a TQT or equivalent cannot be performed. The integrated risk assessment also describes how follow-up studies can be used to understand and predict TdP risk of QTc-prolonging drugs, however, these are evaluated on a case-by-case basis. ECG, electrocardiogram; ICH, International Council for Harmonisation; Q&A, question and answer; QTc, heart rate corrected QT interval; TdP, torsade de pointes.

submitting a complete QT study report, and the database does not capture drugs discontinued as a result of nonclinical hERG and *in vivo* QTc assays. With that said, the 19% number from the FDA database also includes drugs with borderline positive QTc signals (e.g., supratherapeutic dose or high clinical exposure QTc upper bound > 10 ms with QTc mean < 10 ms). Thus, some of the QT study results are false-positive signals for clinical QTc prolongation in patients, and others are true-positive QTc signals, but false-positive signals for TdP risk as not all QTc prolongation leads to TdP, which is discussed further in a later section. Altogether, a pre-test probability of 20–30% seems reasonable.

Regarding the pre-test probability of 10% for TdP, Vargas *et al.*[4] describe that there have been 1,714 total approved drugs and that CredibleMeds.org[27] lists 62 (3.6% of approved drugs) as having a known risk of TdP, whereas there are 240 (14% of approved drugs) classified as known, conditional or potential risk of TdP. The 10%

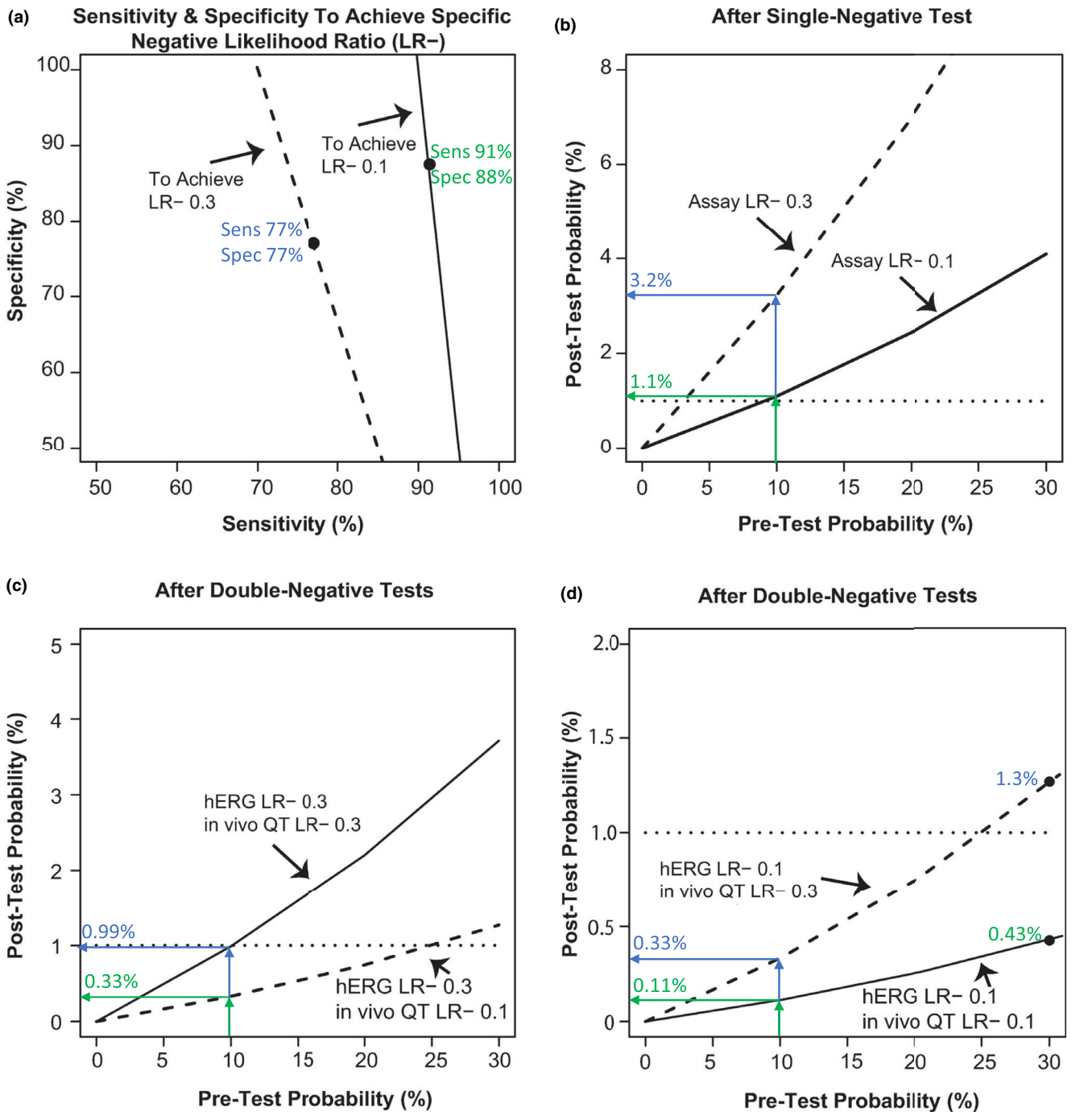
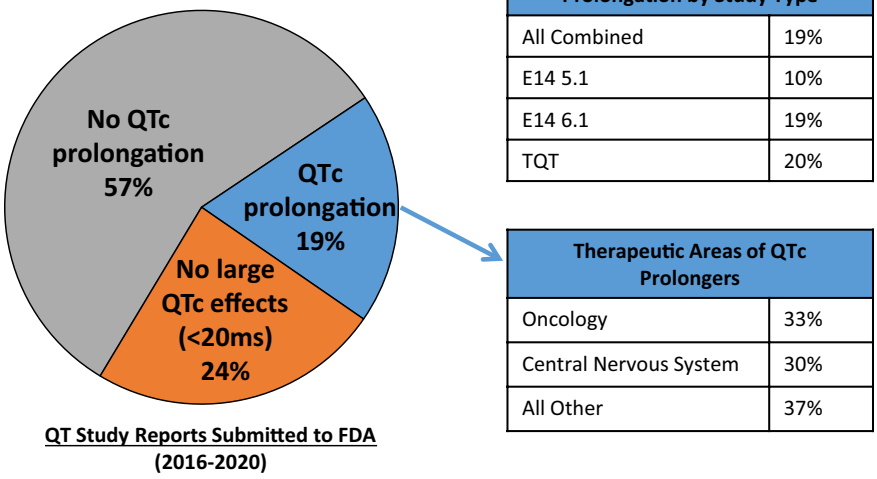


Figure 3 Illustration of how changes in the sensitivity and specificity of an assay on its own (e.g. hERG or *in vivo* QTc alone), or when used in combination with a second assay (e.g. hERG and *in vivo* QTc together), affect the post-test probability of an outcome (e.g. clinical QTc or TdP risk). (**a**) Different sensitivity and specificity combinations for an assay to have moderate (negative likelihood ratio (LR−) = 0.3, dashed line) or high (LR− = 0.1, solid line) discriminatory power when the assay result is negative. As an example, an assay with 77% sensitivity and 77% specificity (dot on dashed line) has moderate discriminatory power (LR− = 0.3), whereas an assay with 91% sensitivity and 88% specificity (dot on solid line) has high discriminatory power (LR− = 0.1). (**b**) Relationship between pre-test and post-test probability after a single negative test. When the pre-test probability is 10% (as used for TdP by Vargas *et al.*), a single negative test with moderate discriminatory power (LR− = 0.3, dotted line) results in a post-test probability of 3.2%, while a single negative test with high discriminatory power (LR− = 0.1, solid line) results in a post-test probability of 1.1% (horizontal dotted line). (**c, d**) Relationship between pre-test and post-test probability for a double-negative drug. Note that a double-negative result where both assays have moderate discriminatory power (LR− = 0.3, solid line in **c**) can bring a pre-test probability of 10% to a post-test probability of 0.99%. When one assay has high discriminatory power (LR− = 0.1), the post-test probability decreases to 0.33% and when both assays have high discriminatory power the post-test probability decreases to 0.11%. As a sensitivity analysis, **d** also shows how even when the pre-test probability is increased to 30%, having one test with moderate and one with high discriminatory power (dot on dashed line) results in a 1.3% post-test probability, whereas having high discriminatory power with both tests (dot on solid line) results in a post-test probability of 0.43%.



Percent of Drugs with QTc Prolongation by Study Type	
All Combined	19%
E14 5.1	10%
E14 6.1	19%
TQT	20%

Therapeutic Areas of QTc Prolongers	
Oncology	33%
Central Nervous System	30%
All Other	37%

Figure 4 The FDA cardiac safety interdisciplinary review team conclusions from clinical QT study reports submitted from 2016 to 2020: 19% were positive for QTc prolongation (i.e., upper bound above 10 ms), 24% were found to have no large QTc effects (i.e., unlikely to have an actual mean QTc effect of 20 ms or larger, primarily drugs under E14 Q&A 6.1 – see Figure 1 legend and text) and 57% were found to have no QTc prolongation (i.e., negative TQT study or equivalent under Q&A 5.1). QTc prolongation was observed in 20% of TQT studies, 10% of E14 Q&A 5.1 studies, and 19% of E14 Q&A 6.1 studies. Considering all study report types together, of the 19% of drugs that prolonged QTc, 33% were for oncology indications, 30% were drugs targeting the central nervous system (i.e., neurology, psychiatry, and anesthesiology/addiction/pain), and 37% were from all other therapeutic indications. These statistics do not capture drugs that underwent clinical QTc evaluation but were discontinued from development prior to the sponsor submitted a clinical QT study report. ECG, electrocardiogram; FDA, US Food and Drug Administration; QTc, heart-rate corrected QT; TdP, torsade de pointes.

number is in between the 3.6% with known TdP risk and the much broader definition of 14%. The broader definition includes conditional TdP risk drugs (n = 51, 3.0% of approved drugs) that are only associated with TdP under certain conditions (e.g., excessive dose, hypokalemia or drug interaction, including when the listed drug is only indirectly associated with TdP by increasing exposure of another TdP-causing drug) and possible TdP risk drugs (n = 138, 8.1% of approved drugs) that do not have evidence of TdP, but can cause QTc prolongation.[27] Although evidence for TdP may develop for some of these drugs, this is also consistent with not all QTc prolonging drugs being associated with TdP. Whereas the implementation of ICH S7B and E14 have likely led to the discontinuation of some drugs that may have otherwise reached the market and caused TdP, it likely took time for these guidelines to have an effect (e.g., drugs being considered for approval in the first 5–7 years after S7B was implemented would have entered clinical development prior to S7B being initiated). Thus, 10% seems reasonable, although a higher number could be used in a sensitivity analysis. **Figure 3** shows how pre-test probability impacts post-test probability.

Sensitivity and specificity of the nonclinical assays for clinical QTc prolongation and TdP liability

Vargas *et al*. combine the data from different studies in the literature, some of which investigate the relationship between hERG or *in vivo* QTc assays and clinical QTc prolongation, and some that investigate the relationship with TdP (see Vargas *et al*. supplementary material).[4] Published data could be subjected to publication bias, and the discriminatory power of each assay is based on an average across studies without considering the number of drugs evaluated in each study. For the hERG assay, there are five studies for each of the endpoints of clinical QTc prolongation and TdP liability, whereas for the *in vivo* QTc assay there are four studies for clinical QTc prolongation and three studies for the *in vivo* QTc assay. Each of the studies includes a varying number of drugs from 12 to 367. The results from each study are taken "as reported" and each study used different methodology, which includes differences in nonclinical assay protocols and the definitions of clinical QTc prolongation or TdP liability.

The largest study for hERG used by Vargas *et al*., which covers clinical QTc and TdP endpoints,[19] includes data from the literature for hERG IC_{50} values, preferentially selecting the most potent value, and for clinical maximum concentration (C_{max}), preferentially selecting the highest exposure, which together means that the data preferentially favors lower hERG safety margins for individual drugs. When evaluating a specific threshold (e.g., hERG margin of 30), this has the effect of increasing sensitivity and decreasing specificity. In general, hERG IC_{50s} (hence safety margins) are sensitive to recording temperature and experimental protocols, including voltage waveforms and stimulation frequencies,[28,29] so it will be critical to have safety margins defined with the same set of protocols used across all drugs. Regarding the *in vivo* QTc assay, our experience from studies submitted to the FDA indicates there is a wide degree of data variability within and among studies and the absence of power analysis to indicate the detectable effect size in each study. Of note, these issues for hERG and *in vivo* QTc assays associated with lack of protocol standardization and lack of explicit statements regarding regulatory expectations are addressed in new S7B Q&As.

Considering these limitations, the Vargas *et al*. data indicates that the two core nonclinical assays have almost equivalent discriminatory power for clinical QTc prolongation (sensitivity 65–67%; specificity 85–86%). Regarding TdP risk, hERG has 88% sensitivity and 81% specificity, whereas *in vivo* QTc has 91% sensitivity and 97% specificity (**Table 1**: "Vargas *et al*. average"). For the limitations discussed above, all these numbers should be interpreted with caution. With regard to the discriminatory power of *in vivo* QTc for TdP risk especially, this should be interpreted with extreme caution as there are only three studies with much smaller sample sizes than for the other assay/endpoint combinations, and the smallest study (12 drugs) increases the average with 100% sensitivity and specificity.

Sensitivity analysis of Vargas *et al*. data for probability of TdP risk

Acknowledging that there are limitations in the underlying datasets, what is the impact of lower sensitivity and specificity on the predicted probabilities for clinical endpoints? To evaluate this, we further assessed the TdP risk endpoint using 10% prior probability of TdP. When using the lowest hERG sensitivity (76%) and specificity (65%) from across any of the studies in the Vargas *et al*. supplement (**Table 1**: "Vargas *et al*. low"), the post-test probability for TdP increases from 1.6% to 3.9%. When using the lowest *in vivo* QTc sensitivity (73%) and specificity (90%) from across any of the studies, post-test probability for TdP increases from 1.0% to 3.2%. In addition, when considering the hERG assay and *in vivo* QTc assay in combination, the post-test probability for TdP increases from 0.15% to 1.2% (**Table 1**). This highlights that while the post-test probability for TdP increases when using the lower sensitivity and specificity, a double-negative result still has a post-test probability of ~ 1%. In addition, if the pre-test probability for TdP is increased by 50% (from 10% to 15%), the post-test probability for TdP increases from 0.15% to 0.23% using the Vargas *et al*. average numbers and from 1.2% to 1.9% using the Vargas *et al*. low numbers. This analysis provides important perspective on how changes in the sensitivity and specificity of individual assays affect the probability of TdP following a double-negative result.

Impact of hERG safety margin threshold

We observed that the hERG assay sensitivity and specificity from the "Vargas *et al*. low" scenario are quite similar to that from a recent study by Ridder *et al*. when a hERG safety margin of 30 is used (**Table 1**).[11] The Ridder *et al*. data come from a 13-site prospective study with a standardized voltage waveform protocol on automated patch clamp systems including the 28 drugs from the Comprehensive *in vitro* Proarrhythmia Assay (CiPA) initiative. It is important to note that the Ridder *et al*. data were collected at ambient temperature, as opposed to physiologic temperature, which is commonly done in Good Laboratory Practice hERG

Table 1 Post-test probability of TdP liability with different hERG assay and *in vivo* QTc assay performance statistics/thresholds with each assay individually and in combination

				Probability for TdP[a] after negative hERG and *in vivo* QTc assays	
hERG assay option	**hERG sensitivity**	**hERG specificity**	**Probability for TdP[a] after negative hERG assay**	**(*in vivo* QTc from Vargas *et al*. average[b])**	**(*in vivo* QTc from Vargas *et al*. low[c])**
hERG safety margin 30 (Vargas *et al*. average[b])	88%	81%	1.6%	0.15%	0.47%
hERG safety margin 30 (Vargas *et al*. low[c])	76%	65%	3.9%	0.38%	1.2%
hERG safety margin 30 (Ridder *et al*.[d])	73%	67%	4.3%	0.42%	1.3%
hERG safety margin 50 (Ridder *et al*.[d])	83%	61%	3.0%	0.29%	0.92%
hERG safety margin 100 (Ridder *et al*.[d])	90%	37%	3.0%	0.28%	0.88%
hERG safety margin 300 (Ridder *et al*.[d])	98%	33%	0.67%	0.06%	0.21%
***In vivo* QTc assay option**	***In vivo* QTc sensitivity**	***In vivo* QTc specificity**	**Probability for TdP[a] after negative *in vivo* QT assay**	-	-
Vargas *et al*. average[b]	91%	97%	1.0%	-	-
Vargas *et al*. low[c]	73%	90%	3.2%	-	-

QTc, corrected QT; TdP, torsade de pointes.
[a]All post-test probabilities for TdP are based on a pre-test (prior) probability of 10%. All numbers in table (sensitivity, specificity, and probability) rounded and presented with two significant digits. [b]Vargas *et al*. sensitivity and specificity numbers come from the main data presented in their paper, which is an average of published studies referenced in their supplement.[4] [c]"Vargas *et al*. low" sensitivity and specificity numbers come from the lowest individual sensitivity and specificity from the referenced studies in their supplement.[4] [d]Ridder *et al*. sensitivity and specificity numbers come from different hERG margin thresholds from that dataset.[11]

studies and is recommended in the new S7B Q&As, and there were significant procedure differences across participating sites that could impact results (e.g., data quality/analysis methods not standardized and drug concentration verification not performed), for which the new S7B Q&As provide best practice recommendations. Thus, we believe the hERG assay performance observed in Ridder *et al*., and the Vargas *et al*. low scenario above, can be improved upon.

The Ridder *et al*. data allow us to probe how changing the hERG safety margin threshold (e.g., from a hERG margin threshold of ≥ 30 to define a negative assay as in Vargas *et al*. to higher margins) affects the post-test probability for TdP. **Table 1** shows how increasing the hERG safety margin threshold increases the sensitivity of the assay but decreases the specificity. When assessing a new drug in development, if a hERG margin threshold of ≥ 30 is used to define a negative assay and the new drug meets the definition for negative, then the post-test probability for TdP is 4.3% (**Table 1**). If the hERG margin threshold is increased to ≥ 50 and the new drug is still negative, then the post-test probability for TdP decreases to 3.0%. However, when using a margin of ≥ 100, the post-test probability of TdP remains at 3.0%. The peculiar result going from a margin of 50 to 100 is achieved because the increase in sensitivity (from 83% to 90%) is offset by a much larger decrease in specificity (from 60% to 36%). This highlights that although the predicted probability of TdP following a negative hERG assay depends more on sensitivity, it is also affected by specificity (see **Figure 3a**, which illustrates how different combinations of sensitivity and specificity can result in the same negative likelihood ratio). This also suggests it does not necessarily make sense to aim for the highest sensitivity if it results in extremely low specificity, especially if two complementary assays (e.g., hERG and *in vivo* QTc) will be used. Indeed, a double-negative finding using the Ridder hERG safety margin of 50 in combination with the Vargas low scenario for *in vivo* QTc results in a TdP probability of < 1% (**Table 1**).

THE ROLE OF DOUBLE-NEGATIVE NONCLINICAL DATA IN THE NEW ICH E14/S7B DRAFT Q&AS (STAGE 1)

The new E14/S7B Q&As recognize the value of double-negative nonclinical data and provide an alternative pathway to assess QTc/TdP liability, reduce the number of TQT studies (revised E14 Q&A 5.1), and inform regulatory decision making (and labeling) at the time of a marketing application when a TQT or equivalent cannot be performed (revised E14 Q&A 6.1).[1]

Revised E14 Q&A 5.1

Under the revised E14 Q&A 5.1,[1] double-negative nonclinical data can be used to allow for additional TQT "substitutes" when the drug exposure in concentration-QTc analysis (from phase I clinical trials) is not high enough to meet the current requirement,[30] which is two times the high clinical exposure. The high clinical exposure is the exposure that the supratherapeutic dose in a TQT study is intended to meet or exceed and defined as the exposure when the mean steady-state maximum concentration associated with the maximum therapeutic dose is increased by the largest intrinsic (e.g., renal or liver impairment) or extrinsic (e.g., drug interaction or food effect) factor.[9] With the revised E14 Q&A 5.1, the clinical data only need to meet the TQT requirement of high clinical exposure when the nonclinical *in vitro* hERG assay and *in vivo* QTc assay are conducted using best practices and both are negative according to definitions summarized below. This has the potential to reduce substantially the number of clinical trials (dedicated TQT studies) in drug development, as from 2016–2020, 32% of QT study reports submitted fell under E14 Q&A 5.1 and only 42% of them covered the 2 times high clinical exposure (**Figure 1**).

Revised E14 Q&A 6.1

Under the revised E14 Q&A 6.1, double-negative nonclinical data can be used to inform regulatory decisions and labeling when a TQT study (or equivalent under E14 Q&A 5.1) cannot be conducted because of safety risks in healthy volunteers (e.g., oncology) or feasibility concerns in patients preclude the use of a positive control or doses to achieve high exposures.[1] This addresses Vargas *et al*.'s second point that the *Conduct of the TQT Study Can be Confounded or Not Feasible*. Under the old Q&A 6.1,[30] when clinical data rule out 10 ms of QTc prolongation but there is no positive control or the data do not meet the Q&A 5.1 requirements, a conclusion of "no large QTc effects" is reached (i.e., the drug is unlikely to have an actual mean QTc effect of 20 ms or larger). Under the revised Q&A 6.1, a conclusion of low likelihood of proarrhythmic effects due to delayed repolarization can be reached. The revised E14 Q&A 6.1 has the potential to have a substantial impact as 24% of QT study reports submitted to the FDA from 2016–2020 fell under E14 Q&A 6.1, and it is difficult to demonstrate lack of QTc effect under the old Q&A leading to unclear risk communication and labeling, such as "no large QTc effects." Under the revised E14 Q&A 6.1, many of these could reach a low-risk determination.

Defining double-negative nonclinical data and additional considerations in the stage 1 Q&As

The new integrated risk assessment Q&As (S7B Q&As 1.1–1.2) present points to consider for when hERG and *in vivo* QTc data are used to support the revised E14 Q&As 5.1 and 6.1 (**Figure 2d**) as summarized below.

- hERG assay – The hERG safety margin should be higher than the safety margin determined based on reference drugs known to cause TdP. The hERG IC_{50} should be determined following S7B Q&A 2.1 best practice considerations, including that the same experimental protocol should be applied to the new drug and the reference drugs. The C_{max} is the mean steady-state maximum plasma concentration when the maximum recommended therapeutic dose is given; considerations regarding clinical vs. high clinical exposure are discussed below. Additional considerations and details are described in the Q&As (S7B Q&As 1.1–1.2 and 2.1).
- *In vivo* QTc assay – 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-specific metabolites should be determined (see ICH S7A Sections 2.3.3.2 and 2.6, and S7B Q&A 3.5). For both E14

Q&As 5.1 and 6.1, the exposure should cover the high clinical exposure scenario. In addition, for E14 Q&A 6.1, the *in vivo* study should have sufficient power to detect a QTc prolongation effect of a magnitude similar to dedicated clinical QT studies. Additional details are described in the S7B best practices considerations Q&As 3.1–3.5.

Regarding the hERG safety margin, an exact threshold is not provided as it can be protocol-dependent.[28,29] Assays following best practices will likely lead to lower thresholds due to reduced variability[19] and historical data for margins have often not considered the high clinical exposure scenario, which will also likely decrease thresholds. In line with this, a recent publication points out when high clinical exposure, including for hERG-blocking metabolites, is considered for some of the intermediate risk drugs for TdP on the CiPA list, their hERG margin decreases significantly.[31] Of note, the Ridder *et al*.[11] data discussed previously is based on the CiPA list of drugs; taking high clinical exposure and the hERG safety margin of metabolites into account will likely result in the sensitivity of the hERG assay increasing compared with what is shown at each hERG margin threshold in **Table 1**. When the hERG safety margin is used to support clinical/regulatory decision making under E14 Q&As 5.1 or 6.1, sponsors will need to provide justification for a given margin according to reference drugs with known TdP risk using the same hERG assay protocol.

Regarding the *in vivo* QTc assay, the new S7B Q&As describe general best practice considerations for study execution and data reporting (Q&As 3.1–3.5) and highlight the need to cover high clinical exposure to support E14 Q&As 5.1 and 6.1. In addition, for E14 Q&A 6.1, the *in vivo* study should have sufficient power to detect a QTc prolongation effect of a magnitude similar to dedicated clinical QT studies. The exact definition of this is not specified in the draft Q&As; however, one could envision that whereas a study would need to be powered to detect the equivalent of a 10 ms human QTc prolongation at the high clinical exposure, if a higher exposure (e.g., 10 times high clinical exposure) is achieved, the study may not need to be powered to detect as small of a QTc prolongation. Additional data may be needed to define this relationship.

ICH E14/S7B WORKING GROUP PROPOSED STAGE 2 OF Q&AS - POTENTIAL EXPANDED ROLE OF DOUBLE-NEGATIVE NONCLINICAL DATA

As outlined in the ICH E14/S7B working group concept paper from 2018 and presentations at the ICH E14/S7B public webinar in 2020, under "stage 2" the working group will consider how to define low-risk drugs that might not require detailed QT-focused clinical evaluation (**Figure 5**).[5,6] As noted by Vargas *et al*.,[4] this is already true for monoclonal antibodies recognized to have a low likelihood of direct hERG channel block-mediated QTc prolongation (existing E14 Q&A 6.3).[30] As discussed during the ICH E14/S7B webinar,[3] initial focus areas for nonclinical data to obviate detailed QT-focused clinical evaluation may include other non-small molecule drugs (e.g., peptides, proteins, and oligonucleotides), drugs with low systemic bioavailability (e.g., dermal or ocular products), and potentially others. Each may require different considerations. Furthermore, experience gained with standardized assays and prospective concordance between double-negative nonclinical data and negative clinical QTc assessments could support further changes.

WHEN HERG BLOCK AND/OR QTC PROLONGATION IS PRESENT IN THE STAGE 1 Q&AS TO ICH E14/S7B

The third and fourth key points raised by Vargas *et al*. regarding how the E14/S7B Q&As can have a meaningful positive impact on drug development are to consider *Multiple Ion Channel Blockade* and *Direct vs. Indirect Effects on QTc Prolongation*.[4] This gets at the critical point that not all types of hERG block and not all types of QTc prolongation are proarrhythmic. This was discussed at the ICH E14/S7B public webinar[3] on the new Q&As and is already described within E14 and S7B, as summarized below.

Previously implemented E14 Q&A 7.1 indicates that the purpose of a TQT study is to characterize the effect of the drug on ventricular repolarization.[30] Furthermore, it states that the TQT study is not intended to assess TdP risk in the target population, but rather to determine whether further data are warranted to assess risk. Additional data can come in the form of more intensive electrocardiogram (ECG) monitoring in late phase clinical trials prior to approval. However, E14 Q&A 7.1 indicates that the recommended intensity of monitoring depends on multiple factors and, in some cases where 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.[30] Specified factors include the magnitude of QTc prolongation and whether it occurs in ordinary use or only when drug concentrations are markedly increased (e.g., renal or hepatic impairment, and drug interactions), other patient characteristics and adverse events in the target population, and other characteristics of the drug, which includes safety pharmacology and pharmacodynamics.[30] This indicates that nonclinical safety pharmacology data and pharmacodynamic data (nonclinical or clinical) could be used as part of an integrated nonclinical-clinical risk assessment. This could include pharmacodynamic data on the exposure-response relationship for QTc and other ECG measurements (e.g., assessing whether the exposure-response relationship for QTc plateaus suggesting lower risk for TdP or assessing ECG biomarkers to differentiate types of multi-ion channel block that have different risk profiles; see the following section on the proposed stage 2 Q&As for additional details). Related to this, the E14 Section 5 on Regulatory Implications, Labelling and Risk Management Strategies,[8] states that:

> Some factors have been proposed that can modify the risk of QT/QTc prolongation. For instance, it has been suggested that some drugs might prolong the QT/QTc interval up to a "plateau" value, above which there is no dose-dependent increase, although this has not been demonstrated adequately to date. It has also been suggested that proarrhythmic risk might be influenced by other pharmacologic effects (e.g., other channel effects).

Proposed Stage 2 of Updates to ICH E14/S7B: How to Define Low-Risk Drugs That Would Not Need Detailed QT-Focused Clinical Evaluation

Monoclonal antibodies do not require detailed QT-focused clinical evaluation

Can we expand to other areas with negative nonclinical data?

- Other non-small molecule products (e.g., peptides, proteins, and oligonucleotides)?
- Drugs with low systemic bioavailability (e.g., dermal or ocular products)?
- Other?

Each may require different considerations and/or nonclinical assays

Figure 5 Proposed stage 2 topic for updating ICH E14/S7B on how to define low-risk drugs that would not need detailed QT-focused clinical evaluation

Thus, ICH E14[8] and its existing Q&As[30] already contains a structure for nonclinical safety pharmacology data and pharmacodynamic data (nonclinical or clinical) to influence phase III clinical trial ECG monitoring and/or labeling for QTc-prolonging drugs. In addition, ICH S7B (Section 2.3.5 - Follow-up Studies)[7] indicates that in circumstances where results among nonclinical studies are inconsistent and/or results of clinical studies differ from those for nonclinical studies, retrospective evaluation and follow-up nonclinical studies can be used to understand the basis for the discrepancies. Results from follow-up studies can be a significant component of an integrated risk assessment.

In the new S7B Q&As, additional details are provided on how follow-up studies (**Figure 2b,c**) can be performed to further explore the mechanisms of QTc prolongation and assess TdP risk as described in the new integrated risk assessment Q&A (**Figure 2d**: S7B Q&A 1.1).[1] If applicable, best practice considerations should be followed for assessment of additional ion channel currents (S7B Q&A 2.1), *in vitro* cardiomyocyte assays (S7B Q&As 2.2–2.4), or *in vivo* studies (S7B Q&As 3.1–3.5). In addition, an appropriately qualified proarrhythmia risk prediction model (S7B Q&As 4.1–4.3) can be used according to its context of use to assess the possibility of TdP in humans. As described earlier, additional details related to these Q&As are available in recent publications.[10–14] In the stage 1 Q&As, while follow-up studies 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, the regulatory impact of the follow-up studies is assessed on a case-by-case basis (**Figure 2d**).[1,32]

ICH E14/S7B WORKING GROUP PROPOSED STAGE 2 OF Q&AS – PROVIDING MORE DETAILED RECOMMENDATIONS FOR QTC-PROLONGING DRUGS

As described above under the double-negative scenario, the ICH E14/S7B working group concept paper from 2018 and presentations at the ICH E14/S7B public webinar in 2020 describe additional proposed "stage 2" updates to E14/S7B (**Figure 6**).[5,6] For stage 2, the working group will consider how nonclinical assays, proarrhythmia models and ECG biomarkers can be used to impact late phase clinical trial design (e.g., intensity of ECG monitoring, patient eligibility criteria, allowed concomitant medications, and patient stopping rules based on QTc prolongation) and regulatory decision making at the time of a marketing application, which informs labeling (**Figure 6d**). This may include a combination of assessing multiple ion channels *in vitro*, proarrhythmia models *in silico* (or *in vitro*, *ex vivo*, or *in vivo*), assays for mechanisms of QTc prolongation beyond direct hERG block, and ECG biomarkers to assess concordance of *in vitro*/*in silico* predictions with *in vivo*/clinical observations (**Figure 6d**). Characterizing the exposure-response relationship is important to determine the potential amount of QTc prolongation in certain patient subgroups that may be subjected to higher drug exposures and evaluate whether the QTc prolongation plateaus, suggesting an indirect mechanism that may be of lower risk. There are multiple potential indirect mechanisms that can lead to QTc prolongation (i.e., not acting through direct ion channel effects), such as from autonomic nervous system effects,[33,34] changes in body temperature,[35] electrolyte abnormalities,[36] and others still being defined (**Figure 6c**).[37] Further investigation should differentiate which mechanisms can be associated with TdP vs. those that are not. Furthermore, an integrated risk assessment could consider a combination of clinical/pharmacodynamic factors with nonclinical assays to rule in or out certain mechanisms if needed.

Regarding multiple ion channel blockade, **Figure 6a** shows how hERG block alone can cause QTc prolongation that can lead to TdP, whereas **Figure 6b** shows how "balanced" multichannel block (late sodium or L-type calcium inhibition in addition to hERG) can lead to QTc prolongation that does not always lead to TdP.[38] This is because hERG block delays repolarization, increasing the probability of L-type calcium current triggering extra beats called early afterdepolarizations (EADs), which can initiate TdP. Because EADs are triggered by inward currents,[39] inhibiting inward currents (late sodium or calcium) can prevent EADs and has antiarrhythmic effects (i.e., not leading to TdP).[40–44] In order to predict TdP risk for a given drug at different exposure levels, ion channel data can be integrated together in a proarrhythmia model (**Figure 6d**). Under CiPA, independent development and validation of an *in silico* proarrhythmia model was performed.[45–48] Following a public workshop in 2018, a white paper on general principles for the validation of proarrhythmia risk prediction

models was developed,[13] which can apply to *in silico* models[45,49,50] or other types of models (*in vitro*,[51] *ex vivo*,[52] and *in vivo*[53]), and these principles are reflected in the stage 1 of Q&As. Independent studies were also performed to test the use of *in vitro* cardiomyocyte assays as independent or complementary assays/proarrhythmia models,[54,55] and best practices for these *in vitro* assays[12] are described in the stage 1 Q&As.[1] Other types of proarrhythmia models can also by qualified under the stage 1 S7B Q&As 4.1-4.3.

Although detailed safety pharmacology assessments can be used under E14 currently to inform the intensity of ECG monitoring for a QTc prolonging drug on a case-by-case basis (E14 Q&A 7.1, discussed previously),[30] the new stage 1 Q&As are expected to usher in more confidence in the nonclinical assays and proarrhythmia models. There still is a need to integrate the nonclinical risk predictions with what is observed clinically (**Figure 6d**). This includes assessing the magnitude of QTc prolongation and whether it occurs with standard use vs. only under certain circumstances. With a QTc prolonging drug that may be low risk due to balanced multi-ion channel effects from *in vitro/in silico* proarrhythmia models, we believe it is important to check for missed or unanticipated effects *in vivo*/clinically. A retrospective analysis of 34 TQT studies[56] followed by three prospective clinical trials[38,57–59] covering 14 drugs/drug combinations identified that the ECG biomarker J-Tpeakc (i.e. early repolarization period from end of the QRS interval (J point) to peak of the T wave (Tpeak), corrected for heart rate) reflected a balance of inward (hERG) and outward (late sodium and/or calcium) current (**Figure 6d**). QTc prolongers with predominant hERG block prolonged J-Tpeakc, but those with balanced multichannel block did not prolong J-Tpeakc.[38,56-59] This biomarker was proposed not to be an independent proarrhythmia

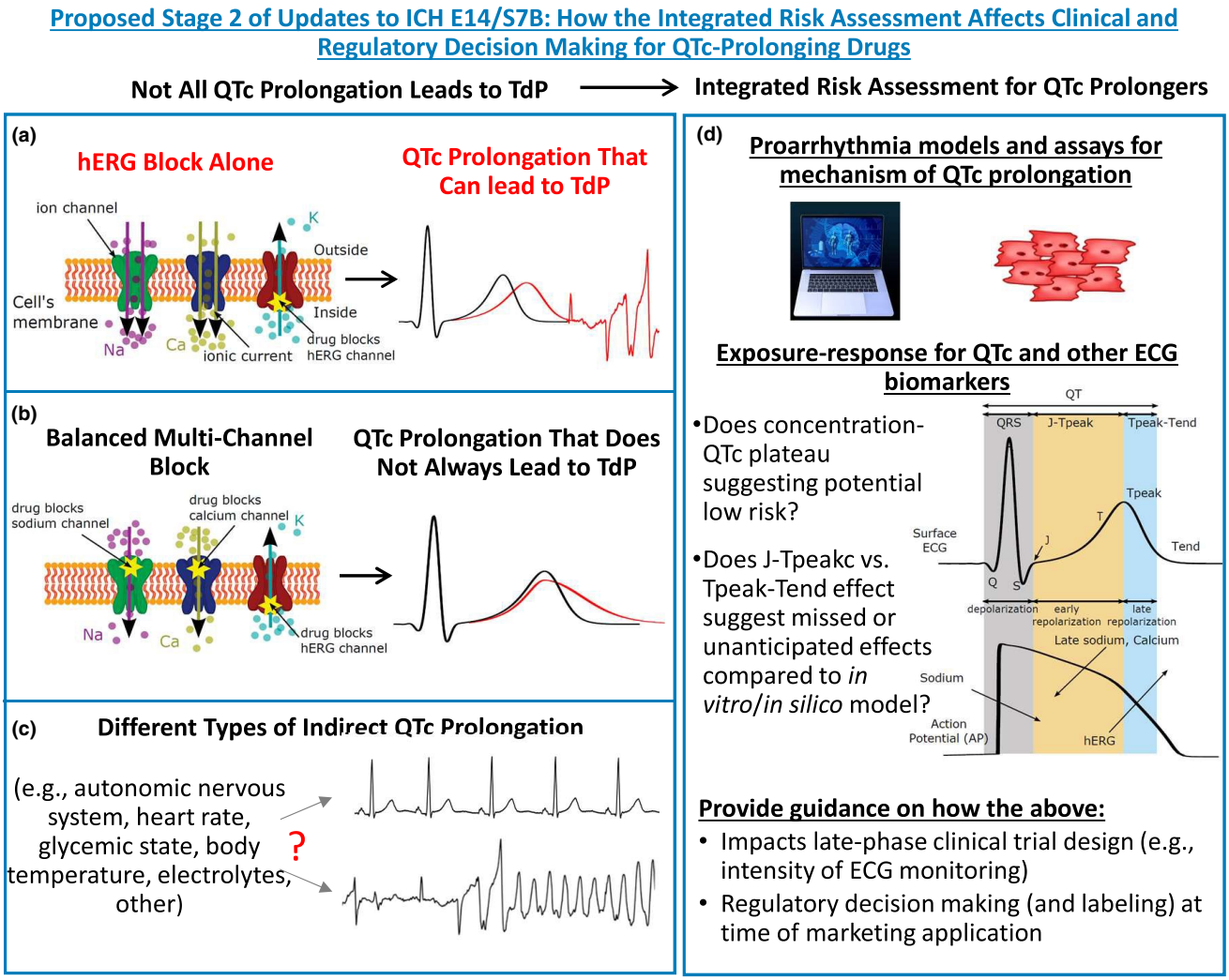


Figure 6 Proposed stage 2 topic for updating ICH E14/S7B -- how an integrated risk assessment including proarrhythmia models and ECG biomarkers can be used to impact clinical and regulatory decision making for QTc-prolonging drugs. (**a**) hERG block alone (or predominant hERG block) causes QTc prolongation that can lead to TdP. (**b**) Balanced multichannel block (hERG block with concomitant inhibition of late sodium or L-type calcium currents) can cause QTc prolongation that does not always lead to TdP. (**c**) Different types of indirect QTc prolongation (i.e., not mediated by directly affecting ion channels) may or may not be associated with TdP. (**d**) Evaluation of TdP risk through an integrated nonclinical-clinical risk assessment. ECG, electrocardiogram; ICH, International Council for Harmonisation; Q&A, question and answer; QTc, heart-rate corrected QT; TdP, torsade de pointes.

marker, but part of an integrated risk assessment checking for missed or unanticipated effects from the *in vitro/in silico* proarrhythmia model.[60,61]

Since the last public workshop on the CiPA initiative in May 2018 and the initiation of the stage 1 Q&As later that year, many additional clinical and nonclinical studies on J-Tpeakc have been published from independent groups. From a review of the literature in November 2020, data on J-Tpeakc now exists for 51 different drugs/interventions, of which there is clinical data for 27 and nonclinical data for 35, with 11 overlapping (see **Supplementary Materials Table S1**). For clinical data, this includes new clinical drug studies,[62–69] comparison of different measurement algorithms,[70–73] improved heart rate assessment correction,[74] and multiple studies from independent groups re-analyzing the FDA clinical trial data.[70,71,75,76] For nonclinical studies, this includes drugs overlapping with clinical study drugs and many interesting drugs without clinical data.[35,77–82] For example, a recent nonclinical *in vivo* study focused on vanoxerine,[80] an interesting drug with large QTc prolongation that has multi-ion channel effects, and was hypothesized by a sponsor to be of relatively low risk for TdP.[83] However, in a phase III trial for atrial fibrillation, 3 of the first 26 patients developed TdP.[84] A nonclinical *in vivo* ECG biomarker study demonstrated that vanoxerine prolonged J-Tpeakc with the same J-Tpeakc vs. Tpeak-Tend pattern as other high-risk drugs (bepridil, sotalol, and E-4031), whereas amiodarone did not prolong J-Tpeakc.[80] This suggests that the J-Tpeakc biomarkerwould have identified vanoxerine as a predominant hERG blocker consistent with high TdP risk.

THE PATH FORWARD

We agree with Vargas *et al.*[4] that the time is here for a fully integrated nonclinical-clinical risk assessment for QTc prolongation and TdP risk. The data they present based on historical *in vitro* hERG and *in vivo* QTc assays suggest that double-negative results support a low risk for clinical QTc prolongation and a lower risk for TdP. Because of the limitations of the data, we caution against citing the exact probability statistics. However, our sensitivity analyses show that when using lower sensitivity and specificity values, a double-negative nonclinical finding can still result in a post-test probability of TdP of ~ 1% (**Table 1**, **Figure 3**). Furthermore, we expect that assays run according to recommendations in the new S7B draft Q&As[1] will have improved performance. The new S7B draft Q&As (**Figure 2**) outline the best practice considerations for the core nonclinical assays and for follow-up studies (including proarrhythmia models) that can be performed if a drug blocks hERG and/or prolongs QTc. Importantly, the S7B integrated risk assessment Q&As and the revised E14 Q&As provide details on how double-negative nonclinical data can be used to reduce the number of clinical TQT studies and inform regulatory decision making when a TQT study or equivalent is confounded or cannot be performed.

An existing E14 Q&A on ECG Monitoring in Late Stage Clinical Trials already describes how safety pharmacology and pharmacodynamic data can influence the intensity of ECG monitoring for QTc prolongers, and the E14 section on Regulatory Implications, Labelling, and Risk Management Strategies recognizes that pharmacodynamic data (plateau in dose/exposure-response relationship and multichannel effects) may influence TdP risk.[30] The proposed stage 2 of updates to E14/S7B (**Figure 6**) will seek to provide additional details on how a fully integrated nonclinical-clinical assessment, including proarrhythmia models and ECG biomarkers, can impact ECG monitoring in late stage clinical trials along with regulatory decision making and labeling for QTc prolonging drugs.[5,6]

The ICH E14/S7B working group will also evaluate how to identify additional low-risk drugs that do not need detailed QT-focused clinical evaluation (**Figure 5**).[5,6] This is already true for monoclonal antibodies[30] and shorter-term goals for stage 2 may include other therapeutic non-small molecule products and drugs with low systemic bioavailability. Vargas *et al.*[4] propose a more ambitious goal of all small chemical molecules that have "double-negative nonclinical findings in a robust, comprehensive and integrated ICH S7B data package." Successful implementation of the initial stages will be critical to see how far we can get with these nonclinical assays or other new translational models and tools (biomarkers) to replace clinical trials for low-risk drugs and advance drugs in development with false-positive signals for proarrhythmia risk with conventional assays/biomarkers. The integrated nonclinical-clinical assessment here can also serve as a model for other safety areas in drug development and regulatory evaluation.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

ACKNOWLEDGMENTS

The authors thank Aanchal Shah, Joel Zirkle, Mohammadreza Samieegohar, and Chase Tarnstrom for their assistance in preparing this manuscript.

FUNDING

No funding was received for this work.

CONFLICTS OF INTEREST

The authors declared no competing interests for this work.

DISCLAIMER

The opinions expressed in this paper are those of the authors and should not be interpreted as the position of the US Food and Drug Administration. As an Associate Editor of Clinical Pharmacology & Therapeutics, David G. Strauss was not involved in the review or decision process for this paper.

1. ICH E14/S7B draft guideline: clinical and nonclinical evaluation of QT/QTc interval prolongation and proarrhythmic potential question and answers <https://database.ich.org/sites/default/files/ICH_E14-S7B_QAs_Step2_2020_0827_0.pdf> (2020). Accessed November 25, 2020.
2. E14 and S7B clinical and nonclinical evaluation of QT/QTc interval prolongation and proarrhythmic potential questions and answers

draft guidance for industry <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e14-and-s7b-clinical-and-nonclinical-evaluation-qtqtc-interval-prolongation-and-proarrhythmic> (2020). Accessed November 25, 2020.
3. New approaches for an integrated nonclinical-clinical QT/proarrhythmic risk assessment <https://www.fda.gov/drugs/news-events-human-drugs/new-approaches-integrated-nonclinical-clinical-qtproarrhythmic-risk-assessment-10152020-10162020> (2020). Accessed November 25, 2020.
4. Vargas, H.M. *et al*. Time for a fully integrated nonclinical-clinical risk assessment to streamline QT prolongation liability determinations: a pharma industry perspective. *Clin. Pharmacol. Ther.* **109**, 310–318 (2021)
5. ICH E14/S7B Implementation Working Group Final Concept Paper <https://database.ich.org/sites/default/files/E14S7B_IWG_Concept_Paper.pdf> (2018). Accessed November 25, 2020.
6. Strauss, D. Background, motivation for and overview of the new Q&As for ICH E14 and S7B. *ICH E14/S7B Webinar on New Approaches for an Integrated Nonclinical-Clinical QT/Proarrhythmic Risk Assessment* <https://sbiaevents.com/files2/ICH-Webinar-Oct-2020.zip> (2020). Accessed November 25, 2020.
7. ICH Harmonised Tripartite Guideline. The non-clinical evaluation of the potential for delayed ventricular repolarization (QT Interval Prolongation) by human pharmaceuticals S7B <https://database.ich.org/sites/default/files/S7B_Guideline.pdf> (2005). Accessed November 25, 2020.
8. ICH Harmonized Tripartite Guideline. The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarryhthmic drugs E14 <https://database.ich.org/sites/default/files/E14_Guideline.pdf> (2005). Accessed November 25, 2020.
9. Garnett, C. Revised E14 Q&As and Presentation of Examples to Highlight the Impact of Nonclinical Data on Clinical Development and Interpretation. *ICH E14/S7B Webinar on New Approaches for an Integrated Nonclinical-Clinical QT/Proarrhythmic Risk Assessment* <https://sbiaevents.com/files2/ICH-Webinar-Oct-2020.zip> (2020). Accessed November 25, 2020.
10. Recommended voltage protocols to study drug-cardiac ion channel interactions using recombinant cell lines <https://www.fda.gov/media/131157/download> (2019). Accessed November 25, 2020.
11. Ridder, B.J. *et al*. A systematic strategy for estimating hERG block potency and its implications in a new cardiac safety paradigm. *Toxicol. Appl. Pharmacol.* **394**, 114961 (2020).
12. Gintant, G. *et al*. Repolarization studies using human stem cell-derived cardiomyocytes: validation studies and best practice recommendations. *Regul. Toxicol. Pharmacol.* **117**, 104756 (2020).
13. Li, Z. *et al*. General principles for the validation of proarrhythmia risk prediction models: an extension of the CiPA in silico strategy. *Clin. Pharmacol. Ther.* **107**, 102–111 (2020).
14. Han, X. *et al*. 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. *J. Pharmacol. Toxicol. Methods* **105**, 106890 (2020).
15. Park, E. *et al*. Can non-clinical repolarization assays predict the results of clinical thorough QT studies? Results from a research consortium. *Br. J. Pharmacol.* **175**, 606–617 (2018).
16. Wallis, R.M. Integrated risk assessment and predictive value to humans of non-clinical repolarization assays. *Br. J. Pharmacol.* **159**, 115–121 (2010).
17. Gintant, G. An evaluation of hERG current assay performance: translating preclinical safety studies to clinical QT prolongation. *Pharmacol. Ther.* **129**, 109–119 (2011).
18. Pollard, C.E. *et al*. An analysis of the relationship between preclinical and clinical QT interval-related data. *Toxicol. Sci.* **159**, 94–101 (2017).
19. Leishman, D.J., Abernathy, M.M. & Wang, E.B. Revisiting the hERG safety margin after 20 years of routine hERG screening. *J. Pharmacol. Toxicol. Methods* **105**, 106900 (2020).
20. Ewart, L. *et al*. The concordance between nonclinical and phase I clinical cardiovascular assessment from a cross-company data sharing initiative. *Toxicol. Sci.* **142**, 427–435 (2014).
21. Webster, R., Leishman, D. & Walker, D. Towards a drug concentration effect relationship for QT prolongation and torsades de pointes. *Curr. Opin. Drug Discov. Devel.* **5**, 116–126 (2002).
22. Redfern, W.S. *et al*. Relationships between preclinical cardiac electrophysiology, clinical QT interval prolongation and torsade de pointes for a broad range of drugs: evidence for a provisional safety margin in drug development. *Cardiovasc. Res.* **58**, 32–45 (2003).
23. Hanson, L.A., Bass, A.S., Gintant, G., Mittelstadt, S., Rampe, D. & Thomas, K. ILSI-HESI cardiovascular safety subcommittee initiative: evaluation of three non-clinical models of QT prolongation. *J. Pharmacol. Toxicol. Methods* **54**, 116–129 (2006).
24. Kramer, J. *et al*. MICE models: superior to the HERG model in predicting Torsade de Pointes. *Sci. Rep* **3**, 2100 (2013).
25. Ando, K. *et al*. QT PRODACT: in vivo QT assay with a conscious monkey for assessment of the potential for drug-induced QT interval prolongation. *J. Pharmacol. Sci.* **99**, 487–500 (2005).
26. Toyoshima, S. *et al*. QT PRODACT: in vivo QT assay in the conscious dog for assessing the potential for QT interval prolongation by human pharmaceuticals. *J. Pharmacol. Sci.* **99**, 459–471 (2005).
27. CredibleMeds QT drugs lists <https://www.crediblemeds.org> (2020). Accessed November 21, 2020.
28. Stork, D. *et al*. State dependent dissociation of HERG channel inhibitors. *Br. J. Pharmacol.* **151**, 1368–1376 (2007).
29. Walker, B.D. *et al*. Inhibition of the human ether-a-go-go-related gene (HERG) potassium channel by cisapride: affinity for open and inactivated states. *Br. J. Pharmacol.* **128**, 444–450 (1999).
30. ICH. E14 guideline the clinical evaluation of QT/QTc interval prolongation and proarrythmic potential for non-antiarrythmic drugs questions & answers (R3) <https://database.ICH..org/sites/default/files/E14_Q%26As_R3_Q%26As.pdf> (2015). Accessed December 3, 2020.
31. Leishman, D.J. Improving prediction of torsadogenic risk in the CiPA in silico model by appropriately accounting for clinical exposure. *J. Pharmacol. Toxicol. Methods* **101**, 106654 (2020).
32. Li, Z. S7B integrated risk assessment Q&As. *ICH E14/S7B Webinar on New Approaches for an Integrated Nonclinical-Clinical QT/Proarrhythmic Risk Assessment* <https://sbiaevents.com/files2/ICH-Webinar-Oct-2020.zip> (2020). Accessed November 25, 2020.
33. Fossa, A.A. *et al*. Dynamic beat-to-beat modeling of the QT-RR interval relationship: analysis of QT prolongation during alterations of autonomic state versus human ether a-go-go-related gene inhibition. *J. Pharmacol. Exp. Ther.* **312**, 1–11 (2005).
34. Harada, T., Abe, J., Shiotani, M., Hamada, Y. & Horii, I. Effect of autonomic nervous function on QT interval in dogs. *J. Toxicol. Sci.* **30**, 229–237 (2005).
35. Boulay, E. *et al*. Confounders and pharmacological characterization when using the QT, JTp, and Tpe intervals in beagle dogs. *Int. J. Toxicol.* **39**, 530–541 (2020).
36. Kim, E.D. *et al*. Associations of serum and dialysate electrolytes with QT interval and prolongation in incident hemodialysis: the Predictors of Arrhythmic and Cardiovascular Risk in End-Stage Renal Disease (PACE) study. *BMC Nephrol.* **20**, 133 (2019).
37. Tran, P.N. *et al*. Mechanisms of QT prolongation by buprenorphine cannot be explained by direct hERG channel block. *PLoS One* **15**, e0241362 (2020).
38. Vicente, J. *et al*. Mechanistic model-informed proarrhythmic risk assessment of drugs: review of the "CiPA" initiative and design of a prospective clinical validation study. *Clin. Pharmacol. Ther.* **103**, 54–66 (2018).
39. January, C.T. & Riddle, J.M. Early afterdepolarizations: mechanism of induction and block. a role for L-type Ca2+ current. *Circ. Res.* **64**, 977–990 (1989).
40. Chezalviel-Guilbert, F., Davy, J.M., Poirier, J.M. & Weissenburger, J. Mexiletine antagonizes effects of sotalol on QT interval duration

and its proarrhythmic effects in a canine model of torsade de pointes. *J. Am. Coll. Cardiol.* **26**, 787–792 (1995).
41. Guo, D., Zhao, X., Wu, Y., Liu, T., Kowey, P.R. & Yan, G.X. L-type calcium current reactivation contributes to arrhythmogenesis associated with action potential triangulation. *J. Cardiovasc. Electrophysiol.* **18**, 196–203 (2007).
42. Badri, M. *et al*. Mexiletine prevents recurrent torsades de pointes in acquired long QT syndrome refractory to conventional measures. *JACC Clin. Electrophysiol.* **1**, 315–322 (2015).
43. Wu, L. *et al*. Role of late sodium current in modulating the proarrhythmic and antiarrhythmic effects of quinidine. *Heart Rhythm* **5**, 1726–1734 (2008).
44. Wu, L. *et al*. Augmentation of late sodium current unmasks the proarrhythmic effects of amiodarone. *Cardiovasc. Res.* **77**, 481–488 (2008).
45. Li, Z. *et al*. Assessment of an in silico mechanistic model for proarrhythmia risk prediction under the CiPA initiative. *Clin. Pharmacol. Ther.* **105**, 466–475 (2019).
46. Dutta, S. *et al*. Optimization of an in silico cardiac cell model for proarrhythmia risk assessment. *Front. Physiol.* **8**, 616 (2017).
47. Chang, K.C. *et al*. Uncertainty quantification reveals the importance of data variability and experimental design considerations for in silico proarrhythmia risk assessment. *Front. Physiol.* **8**, 917 (2017).
48. Li, Z. *et al*. Improving the in silico assessment of proarrhythmia risk by combining hERG (human ether-a-go-go-related gene) channel-drug binding kinetics and multichannel pharmacology. *Circ. Arrhythm. Electrophysiol.* **10**, e004628 (2017).
49. Mirams, G.R. *et al*. Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovasc. Res.* **91**, 53–61 (2011).
50. Tomek, J. *et al*. Development, calibration, and validation of a novel human ventricular myocyte model in health, disease, and drug block. *Elife* **8**, e48890 (2019).
51. Ando, H. *et al*. A new paradigm for drug-induced torsadogenic risk assessment using human iPS cell-derived cardiomyocytes. *J. Pharmacol. Toxicol. Methods* **84**, 111–127 (2017).
52. Champeroux, P. *et al*. Prediction of the risk of Torsade de Pointes using the model of isolated canine Purkinje fibres. *Br. J. Pharmacol.* **144**, 376–385 (2005).
53. Sugiyama, A. Sensitive and reliable proarrhythmia in vivo animal models for predicting drug-induced torsades de pointes in patients with remodelled hearts. *Br. J. Pharmacol.* **154**, 1528–1537 (2008).
54. Blinova, K. *et al*. International multisite study of human-induced pluripotent stem cell-derived cardiomyocytes for drug proarrhythmic potential assessment. *Cell Rep.* **24**, 3582–3592 (2018).
55. Blinova, K. *et al*. Comprehensive translational assessment of human-induced pluripotent stem cell derived cardiomyocytes for evaluating drug-induced arrhythmias. *Toxicol. Sci.* **155**, 234–247 (2017).
56. Johannesen, L. *et al*. Improving the assessment of heart toxicity for all new drugs through translational regulatory science. *Clin. Pharmacol. Ther.* **95**, 501–508 (2014).
57. Vicente, J. *et al*. Assessment of multi-ion channel block in a phase I randomized study design: results of the CiPA phase I ECG biomarker validation study. *Clin. Pharmacol. Ther.* **105**, 943–953 (2019).
58. Johannesen, L. *et al*. Late sodium current block for drug-induced long QT syndrome: results from a prospective clinical trial. *Clin. Pharmacol. Ther.* **99**, 214–223 (2016).
59. Johannesen, L. *et al*. Differentiating drug-induced multichannel block on the electrocardiogram: randomized study of dofetilide, quinidine, ranolazine, and verapamil. *Clin. Pharmacol. Ther.* **96**, 549–558 (2014).
60. Strauss, D.G. *et al*. Comprehensive in vitro proarrhythmia assay (CiPA) update from a Cardiac Safety Research Consortium / Health and Environmental Sciences Institute / FDA meeting. *Ther. Innov. Regul. Sci.* **53**, 519–525 (2019).
61. Vicente, J., Strauss, D.G., Upreti, V.V., Fossler, M.J., Sager, P.T. & Noveck, R. The potential role of the J-Tpeak interval in proarrhythmic cardiac safety: current state of the science from the American College of Clinical Pharmacology and the Cardiac Safety Research Consortium. *J. Clin. Pharmacol.* **59**, 909–914 (2019).
62. Darpo, B. *et al*. Evaluation of the effect of 5 QT-positive drugs on the JTpeak interval—an analysis of ECGs from the IQ-CSRC study. *J. Clin. Pharmacol.* **60**, 125–139 (2020).
63. Täubel, J. *et al*. Effects of the fluoroquinolones moxifloxacin and levofloxacin on the QT subintervals: sex differences in ventricular repolarization. *J. Clin. Pharmacol.* **60**, 400–408 (2020).
64. Matsukura, S. *et al*. Effects of moxifloxacin on the proarrhythmic surrogate markers in healthy Filipino subjects: exposure-response modeling using ECG data of thorough QT/QTc study. *J. Pharmacol. Sci.* **136**, 234–241 (2018).
65. Gal, P. *et al*. First clinical study with AP30663-a KCa2 channel inhibitor in development for conversion of atrial fibrillation. *Clin. Transl. Sci.* **13**, 1336–1344 (2020).
66. Muensterman, E.T. *et al*. Transdermal testosterone attenuates drug-induced lengthening of both early and late ventricular repolarization in older men. *Clin. Pharmacol. Ther.* https://doi.org/10.1002/cpt.2072. [epub ahead of print].
67. Gheorghe, A.C.D. *et al*. Evolution of electrocardiographic repolarization parameters during antiandrogen therapy in patients with prostate cancer and hypogonadism. *Cardiovasc. Toxicol.* **20**, 390–400 (2020).
68. Yiğiner, Ö. *et al*. The effects of supraphysiological oestrogen levels on ventricular repolarisation parameters. *Kardiol. Pol.* 76, 974–979 (2018).
69. Täubel, J. *et al*. The cardiovascular effects of a meal: J-Tpeak and Tpeak-Tend assessment and further insights into the physiological effects. *J. Clin. Pharmacol.* **59**, 799–810 (2019).
70. Brockway, M., Fossa, A.A. & Mason, J.W. Comparison of two highly automated ECG algorithms for detection of drug-induced cardiac ion channel block. *Clin. Pharmacol. Ther.* **104**, 356–363 (2018).
71. Couderc, J.-P. *et al*. An evaluation of multiple algorithms for the measurement of the heart rate corrected JTpeak interval. *J. Electrocardiol.* **50**, 769–775 (2017).
72. Hnatkova, K. *et al*. Detection of T wave peak for serial comparisons of JTp interval. *Front. Physiol.* **10**, 934 (2019).
73. Nunoi Y. *et al*. Electropharmacological analysis of ranolazine in vivo using the halothane-anesthetized dogs. *Proceedings for Annual Meeting of The Japanese Pharmacological Society*. 2020;**93**:3. https://doi.org/10.1254/jpssuppl.93.0_3-o-076
74. Hnatkova, K. *et al*. Heart rate correction of the J-to-Tpeak interval. *Sci. Rep.* **9**, 1–14 (2019).
75. Bystricky, W., Maier, C. & Carter, D. T vector velocity distribution: A new biomarker for identifying drug effects on cardiac myocyte ion channels. *J. Electrocardiol.* **57**, S121 (2019).
76. Bystricky, W., Maier, C., Gintant, G., Bergau, D. & Carter, D. Identification of drug-induced multichannel block and proarrhythmic risk in humans using continuous T vector velocity effect profiles derived from surface electrocardiograms. *Front. Physiol.* **11**, 567383 (2020).
77. Boulay, E. *et al*. A proof-of-concept evaluation of JTPc and Tp-Tec as proarrhythmia biomarkers in preclinical species: a retrospective analysis by an HESI-sponsored consortium. *Int. J. Toxicol.* **38**, 23–32 (2019).
78. Brockway, B., Liddie, S., Moddrelle, D., Morton, X., Delahanty, T. & Hamlin, R. Discrimination of the effects of three cardiac ion channel blockers using ECG biomarkers and arrhythmia incidence in St. Kitts green monkeys (Chlorocebus Sabaeus). *J. Pharmacol. Toxicol. Methods* **99**, 106595 (2019).
79. Wisialowski, T., Gorczyca, W. & Harris, P. Comparison of effects of moxifloxacin on QT, JTpeak and Tpe in conscious telemetered dogs and nonhuman primates. *J. Pharmacol. Toxicol. Methods* **99**, 106595 (2019).
80. Hagiwara-Nagasawa, M. *et al*. Cardiohemodynamic and arrhythmogenic effects of the anti-atrial fibrillatory compound vanoxerine in halothane-anesthetized dogs. *Cardiovasc. Toxicol*.

https://doi.org/10.1007/s12012-020-09612-3. [e-pub ahead of print].
81. Saito, H. *et al*. In vivo comparison of dl-sotalol-induced electrocardiographic responses among halothane anesthesia, isoflurane anesthesia with nitrous oxide, and conscious state. *J. Pharmacol. Sci.* **145**, 16–22 (2021).
82. Matsukura, S. *et al*. Anti-atrial fibrillatory versus proarrhythmic potentials of amiodarone: a new protocol for safety evaluation in vivo. *Cardiovasc. Toxicol.* **17**, 157–162 (2017).