

Virtual Phase I trials predict drug induced QT prolongation from human ventricular tissue

Matteo Costi^{1*} and Jose Felix Rodriguez Matas¹

^{1*}In-Silico Medicine Laboratory, Dept. of Chemistry, Materials and Chemical Engineering “G. Natta”, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milano, 20133, Italy.

*Corresponding author(s). E-mail(s): matteo.costi@polimi.it;

Abstract

Drug-induced QT prolongation remains a central endpoint in early cardiac safety assessment, but Phase I studies are limited in their ability to characterize inter-individual variability and to explore ethically constrained scenarios. We developed an in silico Phase I trial framework that links human-calibrated ventricular cardiomyocyte populations, multichannel ion-channel block, tissue-level propagation, physiological interstitial fibrosis, and pseudo-ECG reconstruction to predict drug-induced QT changes. In the validation compound set, the 4% physiological fibrosis configuration reproduced clinical QT/QTc responses with near-zero bias, a mean absolute error of 2.35 ms, and a root mean square error of 3.10 ms. When inter-patient variability was incorporated through a conservative UB95 endpoint, 60 of 61 validation compounds remained within an ICH-aligned 10-ms upper-bound agreement criterion. Removing physiological fibrosis reduced concordance from 98.4% to 75.4%, indicating that low-degree physiological tissue heterogeneity was a key determinant of predictive performance. These findings support virtual Phase I trials as a human-based digital approach for early QT-focused cardiac safety assessment.

Keywords: Virtual clinical trials, QT prolongation, Cardiac safety, In silico modeling, Drug development

1 Introduction

Cardiac safety assessment remains a critical component of drug development, as drug-induced alterations of ventricular repolarization may lead to clinically relevant QT/QTc prolongation and, in susceptible conditions, life-threatening ventricular arrhythmias. Within this context, the QT/QTc interval has long been used as a clinically measurable marker of delayed ventricular repolarization and potential proarrhythmic liability, and its assessment is embedded in international regulatory guidance for both nonclinical and clinical drug development and International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human use (ICH)[1, 2]. In early clinical development, Phase I studies provide the first controlled opportunity to quantify drug-induced QT/QTc changes in humans under defined exposure conditions. The regulatory framework has progressively evolved from the traditional dedicated thorough QT study toward concentration–QTc and exposure–response analyses performed in early-phase studies, provided that ECG quality, assay sensitivity, and exposure coverage are adequate [2–5]. This evolution has increased the translational relevance of quantitative QT assessment in Phase I, while preserving its role as a regulatory decision-support endpoint.

However, conventional Phase I studies remain costly, resource-intensive [6, 7], and constrained by ethical and logistical limitations. They provide limited access to inter-individual response variability, cannot systematically explore rare or extreme electrophysiological conditions, and may be inappropriate in selected therapeutic areas, such as oncology, where investigational compounds can be intrinsically toxic and direct testing in healthy volunteers may be ethically restricted [8, 9]. In addition, animal-based cardiac safety studies are limited by interspecies differences in ionic currents, repolarization reserve, heart rate, ion-channel expression, and drug sensitivity. These limitations motivate the development of human-based computational strategies able to complement experimental and clinical evidence during early cardiac safety assessment, while also aligning with broader efforts to reduce reliance on animal experimentation where scientifically justified [10–12].

Despite its regulatory relevance, QT prolongation is an imperfect surrogate of proarrhythmic risk. Drugs with different multichannel electrophysiological profiles may produce similar QT changes while carrying different torsadogenic liabilities. This limitation motivated the development of more mechanistically informed paradigms, most notably the Comprehensive in Vitro Proarrhythmia Assay (CiPA), which integrates multichannel pharmacology, in silico electrophysiological modeling, and complementary biomarkers to move beyond the traditional hERG/QT-only paradigm [13–17]. These approaches have substantially advanced the mechanistic interpretation of drug-induced electrophysiological effects, but the direct prediction of clinically measured Phase I QT/QTc outcomes remains a distinct translational challenge.

Computational cardiac electrophysiology has provided a foundation for addressing this challenge [18]. Experimentally calibrated population-of-models frameworks have shown that variability in ionic conductances can reproduce inter-individual differences in cellular electrophysiology and pharmacological susceptibility [19–21]. Human in silico drug trials have demonstrated that virtual ventricular cell populations can outperform animal-based assays in predicting clinical proarrhythmic cardiotoxicity,

and more recent approaches have incorporated demographic and pharmacokinetic covariates to refine risk stratification across patient subpopulations [22, 23]. At higher physiological scales, multiscale and tissue/organ-level simulations have shown that drug effects can be propagated from ion-channel block to pseudo-electrocardiographic biomarkers, including QT changes [24–26]. Tissue- and organ-level models, as well as physics-informed surrogate approaches, have further demonstrated that drug-induced QT changes can be reproduced with encouraging quantitative agreement against clinical observations [27, 28]. More recently, virtual clinical QT exposure–response studies based on three-dimensional whole-heart populations and emulator-based approaches have begun to make such simulations more clinically scalable [29, 30].

Nevertheless, a gap remains between these advances and the practical need for a digital, Phase I-oriented workflow capable of reproducing clinical QT outcomes across a broad pharmacological dataset using a regulatorily interpretable endpoint. Many existing approaches focus on mechanistic biomarkers, such as action potential duration, early afterdepolarizations, qNet-related metrics, or proarrhythmic classification, whereas fewer frameworks directly emulate the Phase I logic of comparing post-dose QT changes with clinical reference values. The explicit integration of human-calibrated virtual patients, drug-specific multichannel ionic block, tissue-level propagation, pseudo-ECG reconstruction, automated QT measurement, and patient-variability-aware agreement assessment remains comparatively limited [28–31].

Here, we present a virtual Phase I trial framework designed to predict drug-induced QT prolongation from human ventricular tissue simulations. The framework links clinically matched free plasma exposure and experimentally derived multichannel pharmacology to human-calibrated virtual ventricular cardiomyocytes embedded in a tissue substrate incorporating electrotonic coupling and physiological interstitial fibrosis. Drug effects are propagated through monodomain tissue simulations, pseudo-ECGs are reconstructed using a fixed virtual electrode configuration, and QT intervals are measured before and after drug administration using an automated tangent-based method. The resulting simulated ΔQT values are directly compared with clinical QT/QTc reference data using both mean-level prediction errors and a conservative patient-variability-aware endpoint based on the upper bound of the 95% of the confidence interval (UB95)[1, 2].

In this setting, the proposed framework was conceived as a digital cardiac-safety workflow rather than as a standalone electrophysiological model. Drug-specific pharmacological information and clinically matched exposure values are transformed into virtual patient-level QT responses, which are then summarized through trial-like endpoints, including mean ΔQT and a conservative UB95-based agreement criterion. Accordingly, the aim of this study was not to classify torsadogenic risk through intermediate cellular biomarkers, but to predict a clinically interpretable QT-based safety endpoint aligned with early Phase I cardiac safety assessment. By evaluating the framework across a broad pharmacological dataset and by testing the contribution of physiological interstitial fibrosis through an ablation analysis, we assessed whether virtual Phase I simulations can reproduce clinical drug-induced QT changes with sufficient quantitative accuracy and robustness to support early translational decision-making.

2 Results

2.1 A digital virtual Phase I workflow for QT-focused cardiac safety

The proposed framework was designed to emulate the main logic of a Phase I QT assessment in a fully computational setting. Drug-specific pharmacological inputs, including clinically matched free plasma exposure and multichannel ion-channel block parameters, were propagated through a cohort of human-calibrated virtual patients embedded in a ventricular tissue substrate. For each virtual patient, the workflow generated matched control and drug-treated pseudo-ECGs, from which patient-level ΔQT values were automatically extracted. These patient-level outputs were then aggregated into compound-level mean responses and conservative UB95-based agreement endpoints for comparison with clinical QT/QTc reference data.

This structure allowed the framework to operate as a digital cardiac-safety assessment pipeline rather than as a single-scale electrophysiological simulation. The final outputs were therefore expressed in the same units and interpretive space as early clinical QT assessment: drug-induced QT change in milliseconds and agreement with a 10-ms upper-bound regulatory criterion.

2.2 Drug dataset and clinical validation cohort

The virtual Phase I workflow was evaluated across a broad pharmacological dataset derived from the Crumb and Kramer ion-channel pharmacology datasets. Overall, 67 compounds were simulated and data are provided in Supplementary Table S3.

Among these, 61 compounds met the predefined eligibility criteria for comparison with clinical QT/QTc reference data and were therefore included in the clinical validation cohort. The remaining six compounds, namely chlorpromazine, diltiazem, nilotinib, lamivudine, metronidazole, and piperacillin, were simulated but excluded from quantitative concordance analysis because they did not satisfy the predefined requirements for Phase I clinical QT/QTc reference data.

The validation cohort included drugs spanning multiple pharmacological classes, including antiarrhythmics, antibacterial agents, antipsychotics, antidepressants, antivirals, antineoplastic/endocrine agents, calcium-channel blockers, CNS-active agents, and other therapeutic categories. The dataset included compounds associated with QT prolongation, QT shortening, and absence of clinically relevant QT prolongation. The complete list of compounds, pharmacological class, regulatory status, pharmacokinetic input parameters, clinical QT/QTc reference values, and in silico predictions is reported in Supplementary Table S3.

2.3 Mean-level prediction of clinical QT changes

The primary validation was performed using the ventricular tissue configuration incorporating 4% physiological interstitial fibrosis. For each compound, the simulated ΔQT at $1 \times EFTPC_{max}$ was summarized as the mean response across the 10 virtual patients and compared with the corresponding clinical QT/QTc change reported in the literature.

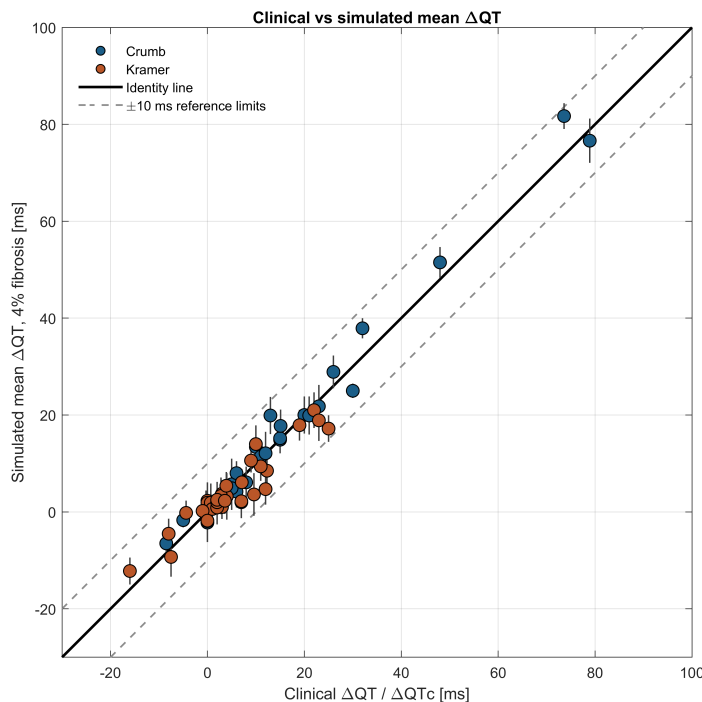


Fig. 1 Mean-level QT prediction accuracy of the 4% physiological fibrosis model. Clinical $\Delta QT/\Delta QT_c$ values are compared with the mean simulated ΔQT across 10 virtual patients. Vertical bars indicate inter-patient standard deviation, the solid line represents identity ($y = x$), and dashed lines denote ± 10 ms visual reference limits around identity. These limits are shown only to aid interpretation of mean-level prediction error and do not replace the UB95-based ICH-aligned agreement criterion. Points are colored by ion-channel pharmacology dataset.

At the mean-response level, the 4% fibrosis model showed strong quantitative agreement with the clinical reference values. The global signed bias was approximately zero (-0.001 ms 95% CI $[-0.78, 0.80]$), indicating the absence of systematic overestimation or underestimation across the validation cohort. The mean absolute error was 2.35 ms 95% CI $[1.85, 2.86]$, while the root mean square error was 3.10 ms 95% CI $[2.47, 3.70]$. The median absolute error was 1.80 ms.

These results indicate that, when the average response of the 10 virtual patients was considered for each compound, the model closely reproduced the QT/QTc prolongation reported in the corresponding clinical studies. The relationship between clinical and simulated ΔQT values is shown in Figure 1. Importantly, because this analysis is based on mean simulated responses, the ± 10 ms lines in Figure 1 should be interpreted only as visual reference limits around the identity line and not as the formal regulatory decision criterion. The conservative regulatory assessment was instead based on the UB95 absolute prediction error, as described in the following subsection.

2.4 Patient-variability-aware UB95 assessment

While the previous analysis assessed mean-level predictive accuracy, we next evaluated whether this agreement was preserved under a more conservative patient-variability-aware regulatory criterion. For each compound, the absolute prediction error was computed for each of the 10 virtual patients as the absolute difference between the patient-specific simulated ΔQT and the corresponding clinical QT/QTc reference value. The upper bound of the 95% confidence interval of the absolute prediction error (UB95) was then calculated and compared with an ICH-aligned 10-ms upper-bound agreement criterion.

Using this UB95-based criterion, 60 out of 61 compounds remained within the 10-ms threshold in the 4% fibrosis configuration, corresponding to an overall conservative ICH-aligned agreement rate of 98.4%. The mean UB95 absolute prediction error was 4.9 ms, with a median value of 4.4 ms. Only ranolazine exceeded the conservative threshold, with an UB95 absolute prediction error of 11.5 ms as shown in Figure 2.

The threshold excess, defined as $\max(\text{UB95} - 10, 0)$, was close to zero in the 4% fibrosis configuration. Specifically, the mean threshold excess was 0.04 ms and the RMS threshold excess was 0.23 ms. These findings indicate that the model remained within the conservative regulatory agreement margin for nearly all compounds even when inter-patient variability was explicitly considered.

To verify that this performance was not driven by a specific pharmacological data source, we performed a source-stratified analysis separating compounds derived from the Crumb and Kramer datasets. UB95-based ICH-aligned agreement rate was preserved in both subsets, with 26/27 Crumb-derived compounds and 34/34 Kramer-derived compounds remaining within the 10-ms threshold in the 4% fibrosis configuration. Therefore, the high overall concordance was preserved despite the use of ion-channel pharmacology parameters derived from different experimental sources (Supplementary Table S4).

Finally, to assess whether the UB95-based agreement performance was disproportionately influenced by a specific therapeutic category, we performed a leave-one-therapeutic-class-out sensitivity analysis. When each therapeutic class was removed in turn, the UB95-based concordance remained stable, ranging from 98.0% to 100.0%. Mean UB95 values also remained within a narrow range (3.16–6.82 ms), with median UB95 values between 2.6 and 6.2 ms. In all leave-one-class-out analyses, the only compound exceeding the 10-ms conservative threshold was ranolazine, except when the “Other cardiovascular” class, which included ranolazine, was removed. These findings support that the conservative UB95-based agreement performance was not driven by any single over-represented therapeutic class (Supplementary Table S5).

2.5 Physiological interstitial fibrosis improves virtual trial performance

To isolate the contribution of physiological interstitial fibrosis an ablation analysis was performed by comparing results with the primary 4% fibrosis configuration with a non-fibrotic configuration. The two models were identical in all other components,

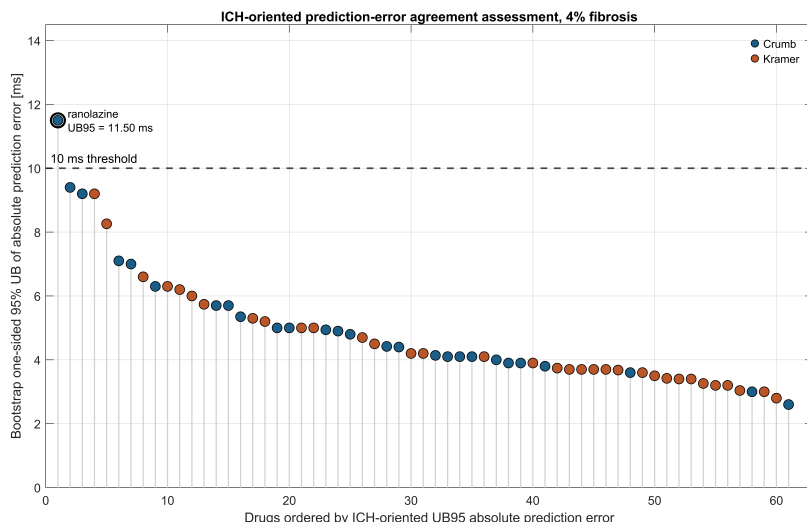


Fig. 2 Conservative UB95-based regulatory assessment of prediction error. For each compound, the UB95 of the absolute prediction error across 10 virtual patients was compared with the 10 ms threshold. Compounds are ordered by decreasing UB95 value; only ranolazine exceeded the threshold in the 4% physiological fibrosis condition. Points are colored by ion-channel pharmacology dataset. Drug identities, UB95 values, threshold-excess values, and corresponding ranking are reported in Supplementary Table S3.

Table 1 Ablation analysis of the contribution of physiological interstitial fibrosis. The 4% fibrosis configuration was compared with an otherwise identical non-fibrotic configuration. Mean-level metrics were computed by comparing the mean simulated ΔQT across the 10 virtual patients with the corresponding clinical QT/QTc reference value. UB95-based metrics were computed from the upper bound of the 95% confidence interval of the absolute prediction error across the 10 virtual patients.

Metric	Non-fibrotic	4% fibrosis
Mean-level predictive accuracy		
Bias [ms]	8.09	-0.001
MAE [ms]	9.97	2.35
RMSE [ms]	19.74	3.10
Median absolute error [ms]	4.10	1.80
Conservative UB95-based regulatory assessment		
Mean UB95 absolute error [ms]	10.3	4.9
Median UB95 absolute error [ms]	4.4	4.4
Compounds with $UB95 \leq 10$ ms	46/61	60/61
Compounds with $UB95 > 10$ ms	15/61	1/61
Conservative concordance [%]	75.4	98.4
Mean threshold excess [ms]	5	0.029
RMS threshold excess [ms]	15.94	0.23
Rescued compounds	-	14
Lost compounds	-	0

including tissue geometry, virtual population, drug-block formulation, pacing protocol, virtual electrode position, pseudo-ECG reconstruction, and QT-measurement algorithm. Therefore, differences between the two configurations reflect the effect of introducing the physiological fibrotic substrate. Importantly, the 4% fibrosis level was selected before the final model-clinical comparison, based on histological evidence supporting a low-degree physiological interstitial fibrotic content in adult myocardium. This value was not fitted on a drug-by-drug basis, was not adjusted to minimize prediction error, and was applied uniformly across all compounds.

Removing physiological fibrosis markedly degraded mean-level predictive accuracy. The non-fibrotic configuration showed a positive bias of 8.09 ms, indicating systematic overestimation of drug-induced QT prolongation. In contrast, the 4% fibrosis configuration reduced the bias to approximately zero. Similarly, MAE decreased from 9.97 ms to 2.35 ms, and RMSE decreased from 19.74 ms to 3.10 ms after introducing physiological fibrosis. Thus, relative to the ablated model, the 4% fibrosis configuration reduced MAE by 7.62 ms and RMSE by 16.64 ms.

This improvement was consistent at the compound level. Absolute prediction error was lower with 4% fibrosis in 49 out of 61 compounds, corresponding to 80.3% of the validation cohort. Paired statistical testing confirmed a significant reduction in absolute prediction error with 4% fibrosis, both with the Wilcoxon signed-rank test ($p = 2.916 \times 10^{-7}$) and with the paired t-test ($p = 6.891 \times 10^{-4}$). Bootstrap analysis further supported this improvement, with a reduction of -7.62 ms in MAE [95% CI: $-12.11, -3.94$] and -16.64 ms in RMSE [95% CI: $-24.42, -7.53$].

The effect of the physiological fibrosis was also evident under the conservative UB95-based regulatory assessment. In the non-fibrotic configuration, 15 of 61 compounds exceeded the 10-ms UB95 agreement margin, whereas only 1 out of 61 compounds exceeded this margin in the 4% fibrosis configuration. Consequently, conservative ICH-aligned agreement rate increased from 75.4% to 98.4% after introducing physiological fibrosis. The mean UB95 decreased from 10.3 ms to 4.9 ms, while the mean threshold excess decreased from 5 ms to 0.029 ms and the RMS threshold excess decreased from 15.94 ms to 0.23 ms. These results, summarized in Table 1, indicate that physiological interstitial fibrosis primarily reduced the large conservative prediction errors that caused threshold exceedance in the non-fibrotic configuration.

The reduction in threshold excess was statistically significant by Wilcoxon signed-rank test ($p = 6.10 \times 10^{-5}$). In the paired threshold-based comparison, 14 compounds were rescued by the inclusion of 4% fibrosis, moving from above to below the 10-ms UB95 agreement margin, while no compound became newly non-concordant. This improvement in threshold-based ICH-aligned agreement rate was significant by exact McNemar/binomial test ($p = 1.22 \times 10^{-4}$).

Ranolazine was the only compound that remained above the conservative UB95 threshold in the 4% fibrosis configuration, with a UB95 absolute prediction error of 11.5 ms. However, even for this compound, physiological fibrosis reduced the threshold excess from 15.8 ms in the non-fibrotic configuration to 1.8 ms in the 4% fibrosis configuration.

Overall, this ablation analysis shows that physiological interstitial fibrosis was a key determinant of the predictive behavior of the framework. Compared with

the non-fibrotic configuration, the 4% fibrosis model reduced systematic overestimation, lowered large compound-specific prediction errors, and substantially increased the number of compounds satisfying the conservative UB95-based 10-ms agreement criterion.

3 Discussion

In this study, we developed and evaluated a digital virtual Phase I trial framework for QT-focused cardiac safety assessment. The framework converts clinically matched drug exposure and experimentally derived multichannel ion-channel pharmacology into patient-level QT responses generated from human-calibrated ventricular tissue simulations. Across 61 validation compounds with available clinical QT/QTc reference data, the framework reproduced drug-induced QT changes with near-zero signed bias, a mean absolute error of 2.35 ms, and a root mean square error of 3.10 ms. When inter-patient variability was incorporated through a conservative UB95-based endpoint, 60 of 61 compounds remained within the 10-ms agreement threshold. These results indicate that virtual Phase I simulations can reproduce clinically observed QT responses with both mean-level accuracy and patient-variability-aware robustness.

A central objective of the present work was to move from purely mechanistic electrophysiological descriptors toward a direct prediction of the endpoint commonly evaluated in early clinical cardiac safety assessment. Previous *in silico* approaches have provided important advances in the prediction of drug-induced electrophysiological effects, including action potential prolongation, repolarization abnormalities, qNet-related metrics, and proarrhythmic classification. However, the direct reproduction of Phase I-like QT outcomes remains challenging because it requires the integration of drug-specific multichannel block, inter-individual electrophysiological variability, tissue-level propagation, pseudo-ECG reconstruction, and a reproducible QT-measurement procedure.

The present framework was therefore designed to preserve this translational chain. Drug action was introduced through experimentally derived multichannel ionic block, applied to a heterogeneous population of calibrated human ventricular cardiomyocytes. These cells were then embedded in a spatially organized ventricular tissue substrate, allowing electrotonic interactions and tissue-level propagation to influence the final pseudo-ECG morphology and QT interval. As a result, the simulated output was not a cellular surrogate of repolarization alone, but a tissue-derived ΔQT that could be directly compared with clinical QT/QTc changes reported in Phase I studies.

The low mean-level error observed in the 4% fibrosis configuration indicates that the model was able to reproduce the central tendency of clinical QT/QTc responses across a heterogeneous pharmacological dataset. Importantly, this agreement was obtained without using clinical QT/QTc data to calibrate model parameters, tune the QT-measurement algorithm, select the virtual electrode position, or optimize the fibrotic configuration. Clinical QT/QTc values were used only as external reference endpoints for validation. This separation between model construction and model evaluation is essential for preserving the interpretability of the validation results.

The primary translational value of the framework is its potential use as a digital decision-support layer in early drug development. In current cardiac safety workflows, compounds are evaluated through a combination of in vitro ion-channel assays, nonclinical models, animal studies, and early clinical evidence. Each of these sources provides useful information, but none of them alone fully captures human inter-individual variability in drug-induced ventricular repolarization. The proposed framework adds a human-based virtual trial layer between experimental pharmacology and clinical QT assessment. In this setting, drug-specific multichannel pharmacology and clinically relevant exposure values are propagated through a cohort of virtual human subjects to generate predicted mean ΔQT , patient-level variability, and a conservative UB95-based agreement endpoint.

This type of output may be useful at different stages of early translational decision-making.

Before first-in-human studies, the framework could support compound prioritization by identifying drugs likely to produce clinically relevant QT prolongation under human-like electrophysiological conditions. During early clinical development, it could help interpret borderline or unexpected QT-related signals by linking multichannel ion-channel block to tissue-level repolarization changes and pseudo-ECG-derived QT prolongation. In addition, because the output is expressed in milliseconds and evaluated against an upper-confidence-bound criterion, it remains directly interpretable within the same conceptual space used in clinical QT/QTc safety assessment.

The framework may be particularly relevant in therapeutic areas where standard Phase I evaluation in healthy volunteers is ethically or practically constrained. This is especially important for selected oncology drugs or other intrinsically toxic compounds, for which administration to healthy volunteers solely for cardiac safety characterization may be inappropriate. In such cases, early clinical evaluation is often performed directly in patients, where disease status, concomitant therapies, altered physiology, and clinical heterogeneity can complicate the interpretation of QT findings. A validated virtual Phase I framework could therefore provide an additional preclinical layer of human-based evidence, supporting cardiac safety assessment before exposing patients to potentially toxic compounds.

Importantly, the proposed framework is not intended to replace clinical evidence or regulatory QT assessment. Rather, it is intended to complement existing evidence streams by providing a reproducible and mechanistically grounded digital estimate of QT liability under human-like electrophysiological conditions. Within this perspective, virtual Phase I simulations may help refine experimental planning, identify compounds requiring closer QT-focused evaluation, and reduce reliance on animal electrophysiology in settings where interspecies translation is known to be limited.

Mean-level agreement provides a useful measure of the central predictive tendency of the model, but it does not fully capture whether the prediction remains reliable when inter-patient variability is considered. This distinction is particularly important in a clinical-trial-like setting, where the interpretation of drug-induced QT/QTc effects depends not only on the average response, but also on the uncertainty surrounding that response. Accordingly, international QT/QTc regulatory guidance emphasizes an upper confidence-bound logic, whereby regulatory concern is associated with the

upper confidence bound of the drug-induced QTc effect exceeding the 10-ms threshold [1, 32–35].

For this reason, the present analysis separated two complementary validation levels. Mean-level metrics, including signed bias, MAE, RMSE, and median absolute error, were used to quantify the quantitative agreement between simulated and clinical ΔQT values. These metrics describe whether the model correctly reproduces the central tendency of the clinical response across compounds. In contrast, ICH-aligned agreement rate was assessed using a conservative patient-variability-aware endpoint, defined as the upper bound of the 95% confidence interval of the patient-level absolute prediction error. The 10-ms threshold was therefore applied exclusively to this UB95-based endpoint, and not to the mean simulated response.

From a digital-trial perspective, this distinction is essential. A virtual trial should not only predict the average response of a compound, but also indicate whether that prediction remains reliable across the simulated cohort. The UB95 endpoint therefore served as a compact patient-variability-aware output that translated virtual patient heterogeneity into a clinically interpretable agreement measure. In this sense, the UB95 criterion allowed the simulated trial to reproduce the upper-confidence-bound logic commonly used in QT/QTc regulatory safety interpretation, while explicitly incorporating inter-patient variability generated by the virtual cohort.

Under this conservative UB95-based criterion, the 4% fibrosis model remained concordant for 60 of the 61 compounds. The mean threshold excess, defined as $\max(\text{UB95} - 10, 0)$, was close to zero, indicating that threshold violations were essentially absent across the validation cohort. This result supports that the model was not only accurate in reproducing the average clinical ΔQT , but also stable when inter-patient variability was explicitly incorporated into the agreement assessment.

The source-stratified analysis further supported the robustness of the framework. UB95-based concordance was preserved in both the Crumb-derived and Kramer-derived subsets, despite the fact that these datasets originate from different experimental pharmacology sources. Therefore, the high overall concordance was preserved despite the use of ion-channel pharmacology parameters derived from different experimental sources. Similarly, the leave-one-therapeutic-class-out analysis showed that the overall performance was not driven by a single over-represented pharmacological category. When each therapeutic class was removed in turn, the conservative concordance remained between 98.0% and 100.0%, and ranolazine remained the only compound exceeding the UB95 threshold except when its own class was removed. These analyses suggest that the observed agreement was not an artifact of a specific data source or therapeutic class composition.

A major finding of this study is that the inclusion of physiological interstitial fibrosis substantially improved the predictive behavior of the framework. The 4% fibrotic configuration was not chosen by optimizing model performance, but prescribed a priori from histological evidence of low-degree interstitial fibrosis in healthy adult myocardium. This is important because the comparison between the non-fibrotic and fibrotic configurations can be interpreted as an ablation analysis rather than as a parameter-fitting exercise.

Removing physiological fibrosis led to a clear degradation in predictive accuracy. The non-fibrotic configuration showed a positive bias of 8.09 ms, indicating systematic overestimation of drug-induced QT prolongation. In contrast, the 4% fibrosis configuration reduced the bias to approximately zero and markedly decreased both MAE and RMSE. The improvement was consistent at the compound level, with lower absolute prediction error in 49 out of 61 compounds. The paired statistical tests and bootstrap confidence intervals confirmed that this improvement was not limited to a few isolated cases.

The same pattern was observed under the conservative UB95-based assessment. In the non-fibrotic configuration, 15 compounds exceeded the 10-ms UB95 agreement margin, whereas only one compound exceeded this threshold after introducing physiological fibrosis. Thus, the 4% fibrosis configuration rescued 14 compounds and did not cause any compound to become newly non-concordant. This finding suggests that the physiological fibrotic substrate mainly corrected large prediction errors produced by the ablated non-fibrotic model.

From a mechanistic perspective, this behavior is consistent with the role of tissue-level electrotonic coupling and fibroblast–myocyte interactions in modulating repolarization dynamics. A purely cardiomyocyte-based tissue may amplify drug-induced repolarization changes in a way that overestimates the QT response observed clinically. The introduction of a low-degree diffuse fibrotic substrate modifies local coupling and tissue-level repolarization patterns, potentially producing a more realistic pseudo-ECG response. However, this interpretation should be considered mechanistic support rather than definitive proof. The present ablation analysis demonstrates that physiological fibrosis is a key determinant of model behavior, but further dedicated analyses would be required to isolate the exact contribution of fibroblast distribution, fibroblast–myocyte coupling strength, and conductivity scaling. Accordingly, the present study should be interpreted as testing the contribution of a physiologically motivated fibrotic substrate within the proposed virtual Phase I workflow, rather than as a formal uncertainty analysis over all possible fibrosis burdens and spatial organizations.

Ranolazine was the only compound exceeding the conservative UB95 agreement criterion in the 4% fibrosis configuration. This result should be interpreted carefully. The threshold violation emerged under the conservative UB95 criterion, where both patient-level variability and the upper confidence bound of the absolute prediction error were considered. The resulting excess over the 10-ms agreement margin was 1.8 ms, indicating a modest residual discordance rather than a large prediction failure.

A plausible explanation for this residual discordance is pharmacological input uncertainty. In the present framework, drug-channel interaction parameters were represented by their reported central values. Specifically, IC_{50} and Hill coefficient values were treated deterministically, and the variability associated with their experimental estimation was not propagated through the electrophysiological model. This simplification may be particularly relevant for ranolazine, whose net electrophysiological effect depends on the balance between multiple ionic currents and may therefore be sensitive to uncertainty in drug-channel interaction parameters [14, 36, 37]. Small variations in

IC_{50} or Hill coefficient values may alter the relative degree of sodium, calcium, and potassium current block and consequently modify the predicted QT response.

Therefore, the ranolazine result should not be interpreted as evidence of a systematic failure of the framework, but rather as an indication that selected compounds with complex multichannel profiles may require explicit pharmacological uncertainty propagation. A formal uncertainty analysis was outside the scope of the present study and should be addressed in future work by sampling IC_{50} , Hill coefficient, protein-binding, and exposure values within experimentally supported ranges. This would allow the framework to generate prediction intervals rather than a single deterministic QT estimate for compounds whose electrophysiological response is highly sensitive to the balance of multichannel block.

Importantly, ranolazine was also the only compound exceeding the UB95 agreement criterion in the leave-one-class-out analyses, except when the class containing ranolazine was removed. This reinforces the interpretation that the residual conservative discordance was compound-specific rather than driven by a broader failure in a pharmacological category or data source.

Overall, several limitations should be acknowledged. First, the simulations were performed in a simplified ventricular wedge rather than in a full anatomical heart model. This choice allowed controlled and computationally tractable tissue-level simulations across a broad pharmacological dataset, but it necessarily limits the ability to reproduce the full complexity of clinical ECG morphology. For this reason, the reconstructed signal was interpreted as a pseudo-ECG, and the primary endpoint was the relative drug-induced QT change rather than the absolute reproduction of a complete clinical ECG.

Second, the virtual Phase I trial included $N = 10$ virtual subjects per compound. This number was intentionally selected to reproduce the scale of a single small dosing cohort within early Phase I clinical trials, while preserving the computational feasibility of tissue-level simulations across a broad pharmacological dataset. Therefore, the resulting UB95 values should be interpreted as patient-variability-aware agreement metrics within this predefined virtual Phase I cohort, rather than as fully converged population estimates.[38, 39].

Future work should quantify the convergence of mean ΔQT , UB95, and agreement rates as larger virtual cohorts become computationally feasible.

Third, the drug-block model was based on a static Hill-type pore-block formulation. This approach is suitable for broad pharmacological screening because it relies on experimentally available IC_{50} and Hill coefficient values, but it does not explicitly describe state-dependent binding, trapping, use dependence, membrane permeation, or dynamic drug-channel interactions [37, 40]. These mechanisms may be relevant for selected compounds, especially those with complex hERG kinetics or multichannel profiles. Incorporating dynamic Markov formulations for specific drug-channel interactions may improve prediction accuracy for compounds with complex electrophysiological behavior.

In the present study, the static Hill-type formulation was selected as a fit-for-purpose approach to simulate QT changes across a broad set of compounds using experimentally available multichannel pharmacology data

Fourth, the present analysis did not propagate uncertainty in IC_{50} and Hill coefficient values. The drug-block parameters were treated deterministically, even though patch-clamp experiments may produce variability due to experimental protocol, cell system, temperature, voltage-clamp conditions, data fitting, and inter-laboratory differences. This limitation is particularly relevant because the ionic block profile is one of the main drivers of the simulated QT response. Future versions of the framework should incorporate pharmacological uncertainty by sampling IC_{50} and Hill coefficient values within experimentally supported ranges.

Fifth, the framework used clinically matched free plasma concentrations derived from available clinical or pharmacokinetic literature. Although this strategy was necessary to simulate exposure levels corresponding to the clinical QT/QTc reference values, it introduces uncertainty related to protein binding, molecular weight conversion, study-specific pharmacokinetics, and differences between reported exposure metrics. More importantly, the present framework did not explicitly model inter-individual metabolic variability. Each compound was simulated at a clinically matched exposure level, but the patient-specific processes determining exposure, including absorption, distribution, metabolism, elimination, and drug-drug interactions, were not represented. This prevents the model from predicting how different patients may reach different free plasma concentrations under the same administered dose.

Sixth, the clinical validation dataset was assembled retrospectively from literature sources. Although eligibility criteria were defined a priori and compounds were not selected based on model performance, the available clinical data were heterogeneous in terms of study design, QT/QTc correction method, exposure reporting, subject characteristics, and dosing regimen. In addition, the present validation used one selected clinical QT/QTc reference per compound rather than integrating multiple clinical studies for the same drug. A more complete retrospective validation would require harmonized clinical datasets with detailed exposure information, subject-level covariates, and consistent QT/QTc analysis methods. Without such information, pooling multiple studies per compound may introduce additional heterogeneity unless the corresponding patient populations and exposure distributions are explicitly modeled.

Last but not least, the number of compounds that can be evaluated by the framework remains limited by the availability of high-quality experimental ion-channel pharmacology data. Broader application will require additional patch-clamp experiments providing reliable IC_{50} and Hill coefficient estimates across the main cardiac ionic currents. Ideally, these experiments should also report uncertainty bounds, protocol details, and replicate-level variability, enabling the propagation of pharmacological uncertainty into QT prediction.

Future developments will focus on three main directions. First, the framework should be extended to incorporate larger virtual populations and explicit demographic or physiological covariates, such as sex, age, disease status, and altered repolarization reserve. This would allow the simulation of subpopulations that are under-represented or difficult to study in early clinical trials.

Second, the pharmacological module should be expanded beyond deterministic static pore block. This includes both the incorporation of dynamic drug-channel interaction models for selected compounds and the propagation of uncertainty in IC_{50} and

Hill coefficient values. A richer pharmacological description would allow the model to evaluate not only one expected QT response, but also the range of plausible QT responses resulting from experimentally observed variability in ion-channel block.

Third, the electrophysiological framework should be coupled with population pharmacokinetic or physiologically based pharmacokinetic models. This represents a key step toward prospective industrial application. By modeling how individual patients metabolize and distribute a compound, a PK/PBPK layer could provide patient-specific free plasma concentration ranges, which could then be passed to the electrophysiological model to predict QT prolongation. Such integration would shift the framework from a retrospective validation tool based on literature-matched exposure values toward a prospective platform capable of simulating patient-specific QT responses from administered dose.

Furthermore, additional experimental patch-clamp data are needed to expand the range of drugs that can be simulated. The predictive capability of the framework depends on the availability of reliable drug-channel interaction parameters. Therefore, systematic generation of IC_{50} and Hill coefficient values across multiple cardiac ion channels would directly increase the number of compounds that can be evaluated and would improve the credibility of future prospective predictions.

Finally, the tissue substrate could be refined by exploring anatomically detailed ventricular geometries, alternative electrode configurations, and uncertainty in physiological fibrosis configurations. Such extensions would help determine which model components are essential for accurate QT prediction and which can remain simplified without compromising translational performance.

To conclude, these findings support the use of virtual Phase I trials as a digital approach for early QT-focused cardiac safety assessment. By integrating human-calibrated electrophysiological variability, clinically matched drug exposure, tissue-level propagation, pseudo-ECG reconstruction, physiological interstitial fibrosis, and a patient-variability-aware agreement endpoint, the framework provides a bridge between mechanistic cardiac modeling and clinical QT interpretation. Future integration with systematic pharmacological uncertainty propagation, patient-specific PK/PBPK modeling, and larger virtual cohorts could further extend the framework from retrospective clinical validation toward prospective digital support for early drug-development decisions.

4 Methods

The proposed framework was implemented as a digital virtual Phase I trial workflow for QT-focused cardiac safety assessment. The workflow was organized into three interconnected layers. The first layer generated a human-calibrated virtual cohort by sampling ventricular cardiomyocyte electrophysiological profiles and assigning them to a tissue substrate. The second layer translated drug-specific clinical exposure and multichannel ion-channel pharmacology into patient-specific tissue simulations and pseudo-ECG signals. The third layer extracted trial-like QT endpoints, including patient-level ΔQT , mean compound-level ΔQT , and a conservative UB95-based agreement metric for comparison with clinical QT/QTc reference data.

Operationally, the workflow consisted of six steps as shown in Figure 3. Namely: generation of virtual patients, construction of the ventricular tissue substrate with physiological interstitial fibrosis, modeling of drug-induced ionic current block, tissue-level electrical propagation, pseudo-ECG reconstruction with automated QT measurement, and clinical endpoint comparison.

4.1 Step 1: Generation of virtual patients

The first step of the proposed framework consisted of generating a cohort of virtual patients to be used for the *in silico* evaluation of drug-induced QT prolongation. This step included two consecutive stages. First, a large population of virtual human ventricular cardiomyocytes was generated and calibrated against experimental data from healthy human donors. Second, the calibrated cellular population was used to construct a cohort of virtual patients by assigning distinct subsets of cells to the nodes of the ventricular tissue mesh. In the context of the virtual Phase I trial, each virtual patient was intended to represent a distinct digital electrophysiological subject rather than an independent isolated-cell simulation. Patient identity was defined by the specific combination of calibrated cardiomyocyte profiles assigned to the tissue mesh and, subsequently, by the patient-specific distribution of physiological fibroblasts. This design allowed the simulated trial to preserve inter-individual variability at the level of both cellular electrophysiology and tissue composition.

4.1.1 Generation and calibration of the virtual cardiomyocyte population

Cellular variability was introduced by varying nine ionic conductances of the ventricular electrophysiological model, namely those associated with I_{CaL} , I_{to} , I_{NaCa} , I_{K1} , I_{Kr} , I_{Ks} , $I_{Na,f}$, $I_{Na,L}$, and I_{NaK} . The corresponding conductance vector was defined as:

$$\mathbf{g} = [g_{CaL}, g_{to}, g_{NaCa}, g_{K1}, g_{Kr}, g_{Ks}, g_{Na,f}, g_{Na,L}, g_{NaK}].$$

Each conductance was independently varied within a range from -30% to $+110\%$ of its baseline value. This interval was selected to introduce a physiologically meaningful degree of ionic heterogeneity, consistently with previous population-of-models approaches [41], while avoiding excessively large conductance perturbations that would increase the probability of generating non-physiological or overtly pathological action potential phenotypes. This choice was motivated by the intended context of use of the present framework, namely the simulation of Phase I-like studies in healthy virtual subjects.

The multidimensional conductance space was sampled using Latin Hypercube Sampling (LHS) [42], which allowed a broad and structured exploration of the nine-dimensional parameter space, consistently with previous *in silico* electrophysiological studies [43]. A total of $Z = 168000$ candidate cardiomyocytes were initially generated. For each sampled conductance combination, the ventricular action potential was simulated using the O’Hara-Rudy human ventricular cardiomyocyte model [44], following the computational strategy previously adopted for the generation of experimentally calibrated ventricular populations [45].

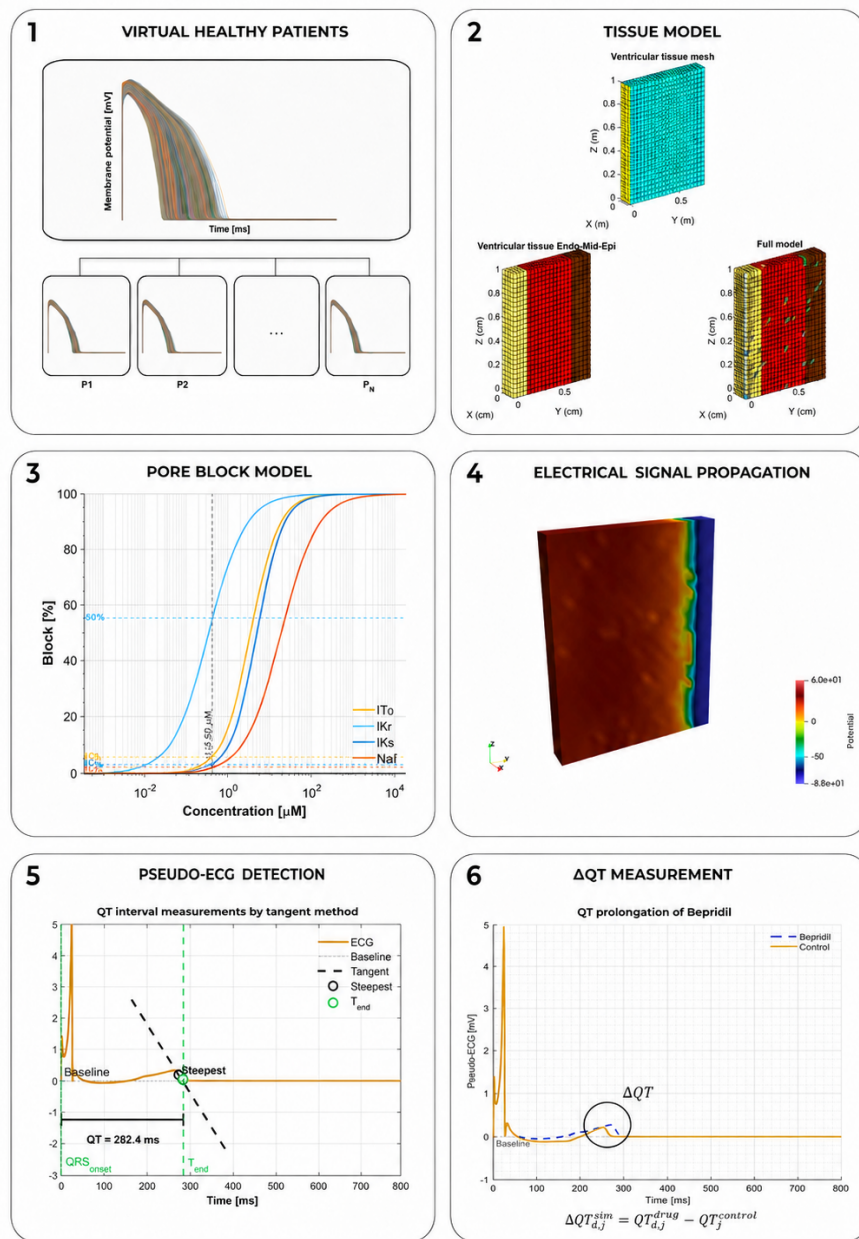


Fig. 3 Workflow for QT-focused cardiac safety assessment. 1) Construction of non-overlapping virtual patients; 2) construction of a heterogeneous ventricular tissue substrate accounting for physiological interstitial fibrosis; 3) pharmacological exposure layer; 4) electrical propagation in the ventricular tissue preparation; 5) digital pseudo-ECG reconstruction and QT-endpoint extraction; 6) virtual trial decision endpoint and comparison with clinical data.

The isolated-cell electrophysiological problem was described as:

$$C_m \frac{dV_m}{dt} = -(I_{tot} + I_{stim}), \quad (1)$$

where V_m is the transmembrane potential, C_m is the membrane capacitance, I_{stim} is the externally applied stimulus current, and I_{tot} is the sum of the ionic membrane currents:

$$\begin{aligned} I_{tot} = & I_{Na} + I_{NaL} + I_{CaL} + I_{to} + I_{K1} + I_{Kr} + I_{Ks} + I_{NaCa} + I_{NaK} \\ & + I_{Cap} + I_{Nab} + I_{Cab} + I_{Kb}. \end{aligned} \quad (2)$$

Here, I_{Na} is the fast sodium current, I_{NaL} is the late sodium current, I_{CaL} is the L-type calcium current, I_{to} is the transient outward potassium current, I_{K1} is the inward rectifier potassium current, I_{Kr} and I_{Ks} are the rapid and slow delayed rectifier potassium currents, respectively, I_{NaCa} is the sodium–calcium exchanger, I_{NaK} is the sodium–potassium pump, I_{Cap} is the sarcolemmal calcium pump, and I_{Nab} , I_{Cab} , and I_{Kb} are background currents.

For each simulated cardiomyocyte, the main action potential biomarkers were extracted and used for calibration against experimental data from healthy human ventricular donors. The calibration biomarkers included resting membrane potential (RMP), action potential amplitude (APA), and action potential duration at 30%, 50%, and 90% of repolarization (APD_{30} , APD_{50} , and APD_{90}) [46]. Only cardiomyocytes whose biomarkers simultaneously fell within the accepted physiological ranges were retained for subsequent analyses.

The acceptance intervals used for calibration were defined according to experimentally derived percentiles and are reported in Table 2. This filtering procedure reduced the initial population from $Z = 168000$ candidate cardiomyocytes to a final pool of $W = 35000$ physiologically calibrated cells. The resulting population represented the pool of healthy human ventricular cardiomyocytes from which virtual patients were subsequently constructed.

Table 2 Calibration percentiles (5th to 95th) used to filter the virtual ventricular cardiomyocyte population [46]

Biomarker	5th percentile	95th percentile
RMP (mV)	-95	-85
APA (mV)	102.7	116.4
APD_{30} (ms)	138	188
APD_{50} (ms)	181	251
APD_{90} (ms)	250	331

4.1.2 Construction of virtual patients

Virtual patients were generated from the calibrated ventricular cardiomyocyte population. Each virtual patient was defined by both an electrophysiological and a structural component. From the electrophysiological point of view, each patient consisted of a specific set of $w = 2184$ calibrated cardiomyocytes. From the structural point of view, these cardiomyocytes were assigned to the nodes of the ventricular tissue mesh, which consisted of 2184 computational nodes. This number of nodes was selected based on sensitivity analyses from previous work showing that smaller sample sizes can adversely affect the reliability of drug-induced proarrhythmic risk assessment [45]. Therefore, the number of cells assigned to each virtual patient was not chosen arbitrarily, but was directly imposed by the spatial discretization of the tissue substrate used for the subsequent monodomain simulations.

To preserve inter-individual variability across the virtual cohort, virtual patients were generated by random subsampling without replacement from the filtered population of $W = 35000$ calibrated cells. In this way, each virtual patient was composed of a disjoint subset of cardiomyocytes, and no cellular profile was shared across different simulated subjects. This non-overlapping sampling strategy was intentionally adopted to avoid artificial replication of nearly identical digital subjects. A substantial overlap in cellular composition between virtual patients would reduce the effective inter-individual variability of the cohort, thereby weakening the ability of the virtual trial to reproduce the heterogeneity typically observed in Phase I clinical studies.

Because virtual patients were generated without replacement, the number of cells per patient and the total number of virtual patients were intrinsically linked. The maximum number of fully disjoint virtual patients that could be generated from the calibrated population was defined as:

$$N_{\max} = \left\lfloor \frac{W}{w} \right\rfloor,$$

where $W = 35000$ is the number of calibrated cardiomyocytes and $w = 2184$ is the number of cells assigned to each patient. Based on this rationale, a cohort of $N = 10$ virtual patients was generated and used for all drug simulations. This cohort size was selected to reproduce, within the proposed virtual Phase I trial framework, the scale of a single small dosing cohort used in early Phase I clinical trials. [38, 39].

After sampling, the selected cardiomyocytes were assigned to the ventricular tissue according to the predefined multilayer organization of the model, which included endocardial, mid-myocardial, and epicardial regions. Endocardial, mid-myocardial, and epicardial phenotypes were not generated as independent calibrated populations. Instead, all cells were sampled from the same experimentally calibrated population and subsequently assigned to the corresponding tissue layers before performing the tissue-level simulations. For each virtual patient, $w = 2184$ calibrated cellular profiles were sampled from the filtered population to match the number of nodes of the ventricular tissue mesh.

4.2 Step 2: Construction of the virtual ventricular tissue substrate

After the generation of the virtual patients, the second step consisted of constructing the ventricular tissue substrate in which each digital subject was embedded. This substrate provided the spatial layer of the virtual Phase I workflow, allowing patient-specific cellular electrophysiological profiles to be mapped onto a controlled ventricular geometry and subsequently propagated through tissue-level simulations.

Each virtual patient was embedded into a cuboid ventricular wedge geometry with dimensions $1.2 \times 8 \times 10 \text{ mm}^3$. The geometry was discretized using a structured hexahedral mesh composed of 1500 linear hexahedral elements and 2184 computational nodes, with a spatial resolution of 0.04 cm along each direction [47]. The mesh therefore provided the spatial substrate onto which the patient-specific cellular electrophysiological properties were mapped, thereby linking the calibrated cellular population to a spatially organized ventricular tissue model.

The ventricular wedge was organized into three transmural regions corresponding to the sub-endocardial, mid-myocardial, and sub-epicardial layers. The relative thicknesses of these layers were defined as 17.84%, 56.34%, and 25.82%, respectively, following the transmural wall-thickening proportions reported by Myers et al. [48]. Layer assignment was performed along the transmural direction before tissue-level simulations. Endocardial, mid-myocardial, and epicardial phenotypes were not generated from independent calibrated populations. Instead, all cardiomyocytes were sampled from the same experimentally calibrated ventricular cell pool and subsequently assigned to the corresponding transmural regions of the tissue.

To reproduce the low-degree structural heterogeneity observed in healthy adult ventricular tissue, physiological interstitial fibrosis was included in the model. Fibrosis was introduced as a fraction of tissue nodes assigned to fibroblasts. Specifically, 4% of the total mesh nodes were randomly selected and replaced by fibroblasts. This percentage was selected according to histological evidence comparing physiological and pathological myocardial fibrosis, supporting a low-degree fibrotic configuration compatible with physiological adult human myocardium [49]. The fibrotic fraction was imposed globally over the entire ventricular wedge and not independently within each transmural layer. As a result, sub-endocardial, mid-myocardial, and sub-epicardial regions were all potentially involved in the random fibroblast distribution.

Within the virtual Phase I workflow, physiological interstitial fibrosis was included to avoid representing healthy virtual subjects as electrically homogeneous cardiomyocyte-only tissues. Instead, each digital subject incorporated a low-degree structural heterogeneity compatible with adult human myocardium. This choice was intended to improve the physiological realism of the simulated tissue response while preserving a controlled and reproducible virtual cohort.

Importantly, the spatial distribution of fibroblasts was generated independently for each virtual patient. Therefore, although the global fibrotic burden was fixed at 4% across the cohort, the specific localization of fibroblast nodes differed from patient to patient. This patient-specific fibroblast distribution was introduced to preserve inter-individual structural variability while maintaining the same physiological level of fibrosis across all virtual subjects.

Fibroblasts were described using the passive fibroblast model proposed by MacCannell et al. [50]. In the present framework, fibroblasts replaced cardiomyocytes at the selected mesh nodes and were electrically coupled to neighboring cardiomyocytes through gap-junctional connections. This implementation was intended to represent physiological interstitial fibroblast–myocyte electrical interaction rather than a pathological scar or replacement fibrosis substrate.

The tissue was modeled as an anisotropic conductive medium. Fiber-dependent propagation was represented using a transversely isotropic formulation of the conductivity tensor. The same baseline anatomical geometry, fiber architecture, and conductivity parameters were used for all virtual patients. This ensured that between-patient differences in the virtual trial arose from calibrated cellular electrophysiological variability and patient-specific physiological fibroblast distribution, rather than from uncontrolled changes in anatomical or numerical settings.

In elements containing only cardiomyocytes, the conductivity was assigned according to physiological values from the literature [47, 51]. In elements containing fibroblasts, the local conductivity was reduced according to a linear scaling rule based on the fibroblast content of the element. Specifically, elements containing only cardiomyocyte nodes retained the physiological cardiomyocyte conductivity, whereas elements containing only fibroblast nodes were assigned one fourth of this value. Intermediate configurations were obtained by linear interpolation according to the number of fibroblast nodes within the element, following an element-wise conductivity scaling rationale previously adopted in computational cardiac electrophysiology studies [52].

The complete implementation details of the fibroblast-dependent conductivity scaling rule are reported in Supplementary Material, Table S1.

Hence, each virtual patient was characterized by a patient-specific distribution of calibrated cardiomyocytes embedded within the same ventricular geometry, multilayer transmural organization, and physiological fibrotic substrate. This strategy allowed the framework to preserve inter-individual electrophysiological and structural variability across virtual patients while maintaining a controlled anatomical and conduction framework.

4.3 Step 3: Pharmacological exposure layer

The third step defined the pharmacological exposure layer of the virtual Phase I trial. For each compound, the simulated exposure was matched to the clinically relevant unbound maximum effective therapeutic plasma concentration, here referred to as $EFTPC_{\max}$. This choice allowed the virtual trial to reproduce the exposure condition associated with the clinical QT/QTc reference endpoint, while maintaining the clinical QT/QTc response fully external to model construction and parameter calibration. The use of unbound concentration was motivated by the fact that the ion-channel block model depends on the free molar concentration available to interact with cardiac ion channels [22, 31, 53].

Whenever available, the free maximum plasma concentration reported in the corresponding clinical QT/QTc study was used directly. If the clinical study reported only

the administered dose, or did not directly provide the associated free plasma concentration, pharmacokinetic data from the literature were used to identify the effective free therapeutic plasma concentration corresponding to the same clinical dosing regimen.

When only total plasma concentration was available, the free concentration was computed as:

$$C_{\text{free}} = f_u C_{\text{tot}}, \quad (3)$$

where C_{tot} is the total plasma concentration and f_u is the unbound fraction of the drug. If plasma protein binding was reported as a fraction P_B , the unbound fraction was computed as:

$$f_u = 1 - P_B. \quad (4)$$

If protein binding was reported as a percentage, P_B was first divided by 100 before applying Eq. 4. When concentrations were reported in mass units, the molecular weight of the compound was used to convert the concentration into molar units before applying the protein-binding correction.

For each drug d , the resulting clinically matched free concentration was defined as:

$$C_d = EFTPC_{\text{max},d}. \quad (5)$$

Thus, each compound was simulated at the free plasma exposure corresponding to the clinical QT/QTc reference used for validation.

Drug-induced ion-channel block was modeled using a static Hill-type pore-block formulation, consistently with previous in silico cardiac safety studies based on ion-channel IC_{50} , Hill coefficient, and therapeutic exposure data [16, 17, 22, 31]. For each drug d and ionic current k , the fractional block was computed as:

$$B_{d,k} = \frac{C_d^{n_{d,k}}}{IC_{50,d,k}^{n_{d,k}} + C_d^{n_{d,k}}} = \frac{1}{1 + \left(\frac{IC_{50,d,k}}{C_d}\right)^{n_{d,k}}}, \quad (6)$$

where $B_{d,k}$ represents the fraction of current k blocked by drug d , $IC_{50,d,k}$ is the half-maximal inhibitory concentration, and $n_{d,k}$ is the Hill coefficient. With this formulation, $B_{d,k} = 0$ indicates no ionic block, whereas $B_{d,k} = 1$ indicates complete block of the corresponding ionic current.

The pharmacological effect was then applied by scaling the corresponding ionic conductance:

$$g_{i,k,d}^{\text{drug}} = g_{i,k}^{\text{control}} (1 - B_{d,k}), \quad (7)$$

where $g_{i,k}^{\text{control}}$ is the baseline conductance of ionic current k in cell i , and $g_{i,k,d}^{\text{drug}}$ is the residual conductance after administration of drug d . Therefore, although the fractional block $B_{d,k}$ was identical for all cells exposed to the same drug, the resulting drug-treated conductance remained cell-specific because it was applied to the heterogeneous conductance values generated during virtual population construction.

The ionic currents considered in the drug-block model were those included in the conductance vector defined in Step 1, namely $ICaL$, I_{to} , $INaCa$, $IK1$, IKr , IKs , INa,f , INa,L , and $INaK$. For each drug, only the currents for which an experimentally reported interaction was available in the adopted pharmacological datasets were

actively blocked [14, 36]. If no experimental block information was available for a given drug-current pair, the interaction was assumed to be absent by setting $IC_{50,d,k} = \infty$ and $n_{d,k} = 1$, which yields $B_{d,k} = 0$.

The IC_{50} and Hill coefficient values were obtained from two experimental ion-channel pharmacology datasets, Crumb et al. [14] and Kramer et al. [36] as indicated in Table S3 in supplementary materials.

These quantities defined the pharmacological input layer of the virtual trial and were used exclusively to simulate drug-induced ionic current block and the resulting QT response. Importantly, the clinical QT/QTc response was not used to fit, select, or adjust any component of the pharmacological model, the QT-measurement algorithm, the virtual electrode configuration, or the physiological fibrosis level. Clinical QT/QTc values were kept fully external to the simulation workflow and were used only after all virtual trial outputs had been generated, as independent endpoints for final comparison.

4.4 Step 4: Modeling of electrical signal propagation

The fourth step consisted of simulating the propagation of the electrical signal through the ventricular tissue substrate. At this stage, the virtual cells generated and pharmacologically modified in the previous steps were no longer treated as isolated cardiomyocytes, but as electrically coupled elements embedded within the ventricular wedge geometry. Electrotonic coupling was introduced to reproduce the functional syncytium of cardiac tissue, in which neighboring cells interact electrically and collectively determine impulse propagation and repolarization dynamics [51].

Electrical propagation was modeled using the monodomain formulation:

$$\nabla \cdot (\mathbf{D} \nabla V_m) = C_m \frac{\partial V_m}{\partial t} + I_{tot} + I_{stim} \quad \text{in } \Omega, \quad (8)$$

with no-flux boundary conditions:

$$\mathbf{n} \cdot (\mathbf{D} \nabla V_m) = 0 \quad \text{on } \partial\Omega. \quad (9)$$

Here, V_m is the transmembrane potential, $\mathbf{D}$ is the effective conductivity tensor, C_m is the membrane capacitance, I_{tot} is the total ionic membrane current, I_{stim} is the externally applied stimulus current, Ω is the ventricular tissue domain, $\partial\Omega$ is its boundary, and $\mathbf{n}$ is the outward unit normal vector.

The monodomain problem was discretized on the structured hexahedral mesh described in Step 2, composed of 1500 linear hexahedral elements and 2184 computational nodes, with a spatial resolution of 0.04 cm. The tissue was modeled as an anisotropic conductive medium using the transversely isotropic conductivity tensor described in Step 2. The model parameters were set as follows: longitudinal conductance $\sigma_L = 0.0023$ S/cm, transverse-to-longitudinal conductance ratio $r = \sigma_T/\sigma_L = 0.35$, and membrane capacitance $C_m = 1.0$ $\mu\text{F}/\text{cm}^2$. The wedge was treated as a transversely isotropic medium, with the conductivity tensor defined as:

$$\mathbf{D} = \sigma_L \left[(1 - r) \mathbf{n}_f \otimes \mathbf{n}_f + r \mathbf{I} \right] \quad (10)$$

where $\mathbf{n}_f$ denotes the local fiber direction and $\mathbf{I}$ is the identity tensor.

Monodomain simulations were performed using an implicit solver. The temporal integration time step was fixed at 0.02 ms for all virtual patients, compounds, and control simulations.

For each virtual patient and drug condition, the ionic properties assigned to the mesh nodes were those resulting from the calibrated cellular population and the drug-induced conductance scaling described in Step 3. Cardiomyocyte nodes were described using the O’Hara–Rudy ventricular ionic model, whereas fibroblast nodes were described using the passive MacCannell fibroblast model. Fibroblast-containing regions contributed to propagation through the local conductivity scaling rule described in Step 2 and through their electrotonic coupling with neighboring cardiomyocytes.

Simulations were performed using the ELVIRA software [47]. The tissue was stimulated at one end of the ventricular wedge, allowing the electrical wavefront to propagate along the main fiber direction. A rectangular stimulus of 2 ms duration was applied with an amplitude set to twice the stimulation threshold. Subsequently, a train of 500 stimuli was delivered at a frequency corresponding to 75 beats/min, with a Basic Cycle Length (BCL) of 800 ms, reproducing a physiological sinus rhythm under controlled pacing conditions.

The same pacing protocol, mesh, fiber architecture, and baseline conductivity parameters were used for all virtual patients and drug conditions. In this way, differences in simulated electrical propagation and QT prolongation arose from the calibrated inter-individual electrophysiological variability, the patient-specific distribution of physiological fibroblasts, and the drug-specific ionic current block, rather than from changes in the imposed stimulation protocol or numerical setup.

4.5 Step 5: Digital pseudo-ECG reconstruction and QT-endpoint extraction

The fifth step converted tissue-level electrophysiological simulations into a digital QT endpoint. For each virtual patient and drug condition, a pseudo-ECG was reconstructed from the simulated transmembrane potentials and processed through the same automated QT-measurement pipeline. The purpose of this step was not to reproduce the full morphology of a clinical 12-lead ECG, but to generate a standardized ventricular repolarization signal from which relative drug-induced QT changes could be measured reproducibly across all virtual patients and compounds.

Since the present simulations focused exclusively on ventricular tissue and did not include atrial chambers, the reconstructed signal cannot be interpreted as a complete clinical ECG. In particular, the P wave and all atrial depolarization and repolarization components were outside the scope of the model. For this reason, the reconstructed signal is referred to as a pseudo-ECG [54, 55].

The pseudo-ECG was computed using a Geselowitz-based formulation:

$$V_o(r) = \frac{1}{8\pi} \int_{\Omega} \nabla V_m(r') \cdot \nabla \left(\frac{1}{|r - r'|} \right) dr', \quad (11)$$

where $V_o(r)$ is the extracellular potential evaluated at the virtual electrode position r , $V_m(r')$ is the transmembrane potential within the ventricular tissue domain Ω , and r' denotes the spatial integration variable over the tissue volume. The integral was computed over the entire ventricular wedge at each time increment [54].

The factor $\frac{1}{8\pi}$ was used as a fixed scaling and sign convention in the pseudo-ECG reconstruction. Since the QT interval was extracted from temporal features of the signal, rather than from its absolute amplitude, this constant affects the amplitude and polarity of the reconstructed pseudo-ECG but does not alter the temporal comparison between control and drug-treated simulations, provided that the same convention is used throughout the entire dataset.

The virtual electrode was placed above the center of the upper surface of the ventricular wedge. Specifically, its position was defined at the midpoint of the x - and z -directions, while its y -coordinate was shifted by 0.003 cm outward from the tissue surface.

This electrode position was kept identical for all virtual patients, drugs, and concentration levels to ensure reproducibility and comparability across simulations. Although the absolute pseudo-ECG morphology may depend on electrode location, the present study evaluated drug-induced QT prolongation as a relative within-patient change between post-administration and control conditions:

$$\Delta QT = QT_{\text{drug}} - QT_{\text{control}}.$$

This fixed-lead strategy preserved internal consistency within the virtual trial. Because each ΔQT value was computed as a within-patient difference between drug-treated and control simulations, the predicted QT change reflected a drug-induced temporal shift in ventricular repolarization rather than differences in virtual electrode placement or signal acquisition.

QT-endpoint extraction was performed using a tangent-based approach for the identification of the end of ventricular repolarization, consistent with established recommendations for QT measurement and interpretation [56–58].

4.6 Step 6: Clinical comparison and virtual trial decision endpoints

The sixth step defined the clinical comparison and decision layer of the virtual Phase I trial. For each compound, patient-level simulated QT changes were aggregated into trial-like endpoints and compared with the corresponding clinical QT/QTc reference value. This step translated the mechanistic simulation outputs into quantities directly interpretable in early cardiac safety assessment, including patient-level ΔQT , compound-level mean ΔQT , patient-level absolute prediction error, and a conservative UB95-based agreement endpoint.

For each virtual patient, the QT interval was measured under control conditions and after drug administration using the tangent-based procedure described in Step 5. Drug-induced QT prolongation was then computed as:

$$\Delta QT_{d,j}^{\text{sim}} = QT_{d,j}^{\text{drug}} - QT_j^{\text{control}}, \quad (12)$$

where $\Delta QT_{d,j}^{\text{sim}}$ is the simulated QT change for drug d in virtual patient j , $QT_{d,j}^{\text{drug}}$ is the QT interval after administration of drug d , and QT_j^{control} is the corresponding control QT interval in the same virtual patient.

For each compound, the final mean-level in silico prediction was reported as the mean and standard deviation of $\Delta QT_{d,j}^{\text{sim}}$ across the $N = 10$ virtual patients.

The standard deviation across virtual patients was used to quantify inter-individual variability within the virtual cohort. Thus, each compound was associated with both an average predicted QT response and a virtual patient-level variability profile.

4.6.1 Eligibility criteria for quantitative validation

Clinical reference data were selected before performing the model-trial comparison. A compound was considered eligible for quantitative validation if it satisfied three predefined criteria: (i) experimental ion-channel pharmacology parameters were available from the Crumb or Kramer datasets, allowing drug-specific pore-block simulation; (ii) a quantifiable clinical QT/QTc change was reported in the literature; and (iii) the clinical QT/QTc reference was obtained from Phase I clinical studies conducted in healthy subjects or in subjects without reported cardiac disease, consistently with the intended context of use of the present framework.

Compounds for which ion-channel pharmacology parameters were available but no quantifiable QT/QTc clinical reference could be retrieved, or for which the available QT/QTc evidence was not obtained from healthy subjects or subjects without reported cardiac disease, were not included in the quantitative concordance analysis. This exclusion criterion was based exclusively on the predefined eligibility criteria and not on model performance.

Clinical QT/QTc data were not used as model inputs and were kept fully external to the virtual trial workflow until the final endpoint-comparison stage. All compounds satisfying the eligibility criteria for quantitative comparison are reported in Supplementary Table S3. Compounds marked as N/A were simulated but excluded from the quantitative validation because they did not meet the predefined eligibility criteria.

4.6.2 Clinical endpoint extraction

The clinical endpoint was extracted as the reported mean drug-induced QT or QTc change, according to the terminology used in the original source. Since all in silico simulations were performed at a fixed Basic Cycle Length of 800 ms, corresponding to 75 beats/min, no heart-rate correction was required for the simulated pseudo-ECG. Therefore, the simulated endpoint is reported as ΔQT . When the clinical source explicitly reported QT or QTc, this information was retained in the dataset; otherwise, the endpoint was retained as reported and used as the available clinical reference.

Clinical reference values included positive QT changes, negative QT changes, or absence of clinically relevant QT prolongation. The sign of the reported QT/QTc change was retained whenever available.

All compounds included in the quantitative validation dataset were selected a priori, before evaluating the agreement between simulated and clinical QT/QTc changes. No compound was added, removed, or reclassified on the basis of model performance. Likewise, clinical QT/QTc data were not used to calibrate model parameters, adjust the QT-measurement algorithm, tune the pore-block formulation, or select the 4% fibrosis level and virtual electrode position. They were used only as independent reference endpoints for comparison with the *in silico* predictions.

4.6.3 Mean-level prediction error and UB95-based agreement endpoint

For each compound d , the simulated QT response was available for $N = 10$ virtual patients. Mean-level predictive accuracy was assessed by comparing the mean simulated response with the corresponding clinical QT/QTc reference value:

$$e_d = \overline{\Delta QT_d^{\text{sim}}} - \Delta QT_d^{\text{clin}}, \quad (13)$$

where $\overline{\Delta QT_d^{\text{sim}}}$ is the mean simulated ΔQT across the 10 virtual patients and $\Delta QT_d^{\text{clin}}$ is the clinical QT/QTc change reported for compound d . Mean-level predictive accuracy was quantified using signed bias, mean absolute error (MAE), root mean square error (RMSE), and median absolute error computed from e_d .

To account for inter-patient variability, prediction error was also quantified at the patient level:

$$AE_{d,j} = |\Delta QT_{d,j}^{\text{sim}} - \Delta QT_d^{\text{clin}}|, \quad j = 1, \dots, N. \quad (14)$$

For each compound, the upper bound of the 95% confidence interval of the mean patient-level absolute prediction error was denoted as $UB95_d$. This quantity was used as a conservative patient-variability-aware agreement endpoint. A compound was considered concordant with the 10 ms agreement criterion if:

$$UB95_d \leq 10 \text{ ms}. \quad (15)$$

Thus, the UB95 endpoint was used as a digital analogue of an upper-confidence-bound decision criterion, allowing the virtual trial to account not only for the average predicted response but also for the uncertainty introduced by inter-patient variability within the simulated cohort.

The 10 ms threshold was selected according to the upper-confidence-bound logic used in QT/QTc cardiac safety assessment [1, 32, 33, 35]. In the present framework, this criterion was applied to the conservative UB95 absolute prediction error, whereas mean-level signed error, MAE, RMSE, and median absolute error were used to quantify predictive accuracy.

4.7 Assessment of the contribution of physiological interstitial fibrosis

To evaluate the role of physiological interstitial fibrosis in the proposed in silico Phase I framework, an additional analysis was performed by comparing the primary tissue configuration, including 4% physiological fibrosis, with an otherwise identical non-fibrotic configuration. The 4% fibrotic burden was selected a priori based on histological evidence of low-degree interstitial fibrosis in adult myocardium and was not optimized to improve agreement with clinical QT data.

In the non-fibrotic configuration, all tissue nodes were assigned cardiomyocyte phenotypes sampled from the experimentally calibrated virtual cell population.

All other model components were kept unchanged between the two configurations, including the ventricular geometry, mesh resolution, transmural layer organization, drug-block formulation, pacing protocol, virtual electrode position, pseudo-ECG reconstruction, and QT-measurement algorithm. Therefore, differences between the 0% and 4% configurations can be attributed to the inclusion of physiological interstitial fibrosis rather than to changes in numerical or pharmacological settings.

For each drug, the in silico ΔQT was computed in both configurations at $1 \times EFTPC_{\max}$, corresponding to the primary exposure level used for comparison with clinical trial data. The effect of physiological fibrosis was quantified by comparing the mean predicted ΔQT , the absolute prediction error with respect to the clinical reference, and the number of compounds classified as clinically concordant according to the 10-ms agreement criterion. This analysis was intended to assess the methodological contribution of including a physiologically fibrotic substrate, and not to perform a formal uncertainty quantification over fibrosis burden.

4.8 Statistical analysis

Mean-level predictive performance was quantified across the 61 validation compounds using signed bias, mean absolute error (MAE), root mean square error (RMSE), median absolute error, and the percentage of compounds whose mean-level absolute prediction error was within 10 ms of the corresponding clinical QT/QTc reference value.

To incorporate the dispersion of prediction errors across virtual patients into the assessment of model-to-clinical agreement, an ICH-oriented UB95-based agreement analysis was performed. For each compound, virtual-patient-level absolute prediction errors $AE_{d,j}$ were computed across the $N = 10$ virtual patients, as defined in Eq. 14. The upper bound of the one-sided 95% confidence interval of the mean absolute prediction error was computed as:

$$UB95_d = \overline{AE}_d + t_{0.95, N-1} \frac{s_{AE,d}}{\sqrt{N}}, \quad (16)$$

where $\overline{AE}_d$ and $s_{AE,d}$ are the mean and sample standard deviation of $AE_{d,j}$ across virtual patients, respectively, and $t_{0.95, N-1}$ is the 95th percentile of the Student's t -distribution with $N - 1$ degrees of freedom. A compound was considered predictively

concordant if:

$$UB95_d \leq 10 \text{ ms.} \quad (17)$$

To quantify the magnitude of threshold violation, the UB95-based prediction-error threshold excess was defined as:

$$E_d^{\text{reg}} = \max(UB95_d - 10, 0). \quad (18)$$

Thus, $E_d^{\text{reg}} = 0$ for compounds satisfying the one-sided UB95-based 10-ms prediction-error criterion and becomes positive only when the upper bound of the one-sided 95% confidence interval of the mean absolute prediction error exceeds this threshold.

The 10-ms criterion was selected to align the model-validation endpoint with the upper-confidence-bound rationale used in ICH E14 QT/QTc assessment. In the ICH E14 framework, regulatory attention is linked to whether the upper bound of the one-sided 95% confidence interval around the estimated drug-induced QTc effect excludes a 10-ms effect; equivalently, this criterion may be expressed using the upper bound of a two-sided 90% confidence interval [1, 32, 35]. In the present study, this rationale was adapted to validation of the virtual Phase I framework by applying the same 10-ms upper-bound criterion to the mean absolute prediction error between simulated ΔQT values and the corresponding clinical QT/QTc reference value. Accordingly, $UB95_d$ should be interpreted as an ICH-oriented model-agreement endpoint rather than as a direct estimate of drug-induced clinical QTc effect or as a formal regulatory classification of the compound.

The 95% confidence intervals reported for global performance metrics were obtained by non-parametric bootstrap resampling across the 61 validation compounds. Compounds were resampled with replacement 5000 times, while preserving the paired correspondence between the non-fibrotic and 4% fibrosis configurations. For each bootstrap replicate, global performance metrics were recomputed, and 95% confidence intervals were defined as the 2.5th and 97.5th percentiles of the resulting bootstrap distributions.

Paired comparisons between the non-fibrotic and 4% fibrosis configurations were performed across the same 61 compounds. Differences in mean-level absolute prediction error and UB95-based threshold excess were assessed using the Wilcoxon signed-rank test. Paired t -tests were additionally computed as parametric sensitivity analyses.

For the paired threshold-based UB95 analysis, compounds were classified as rescued if they exceeded the 10-ms one-sided UB95 threshold in the non-fibrotic configuration but not in the 4% fibrosis configuration, and as lost if they were within the threshold in the non-fibrotic configuration but exceeded it in the 4% fibrosis configuration. The imbalance between rescued and lost compounds was evaluated using an exact McNemar/binomial test on discordant pairs.

To assess whether model performance was influenced by the source of ion-channel pharmacology data, a source-stratified analysis was performed by separately summarizing the Crumb-derived and Kramer-derived compounds. To assess whether the UB95-based model-agreement performance was disproportionately driven by any single therapeutic category, a leave-one-therapeutic-class-out sensitivity analysis was

performed. In this analysis, each therapeutic class was removed in turn, and the UB95-based concordance, mean UB95, median UB95, and threshold excess were recomputed on the remaining compounds.

Data availability. The clinical QT/QTc reference values, pharmacokinetic input parameters, ion-channel pharmacology sources, simulated QT responses, prediction errors, UB95 values, and ICH-aligned agreement rate metrics used in this study are provided in the Supplementary Information. Additional simulation outputs generated during the current study are available from the corresponding author upon reasonable request.

Code availability. The MATLAB code used in this study represents an implementation of the computational methodology described in the Methods and Supplementary Information and is not publicly available because it is subject to intellectual-property protection. All input data, clinical QT/QTc reference values, pharmacological parameters, simulated QT responses, prediction errors, UB95 values, and ICH-aligned agreement rate metrics required to interpret and verify the reported results are provided in the Supplementary Information. The code may be made available to editors and reviewers upon request, subject to institutional and intellectual-property restrictions.

Acknowledgements. The authors acknowledge In-Silico Medicine Laboratory, Department of Chemistry, Materials and Chemical Engineering “G. Natta”, Politecnico di Milano, for providing the scientific environment in which this work was developed.

Ethics approval and consent to participate. Not applicable. This study was entirely computational and did not involve the recruitment of human participants, the conduct of clinical trials, the collection of human biological material, or the performance of new animal experiments. Human- and animal-derived data used in this work were obtained exclusively from previously published literature sources.

Funding. Funded by the European Union – NextGenerationEU. CUP: D43C22002720001.

Author contributions. M.C. and J.F.R.M. conceptualized the study. M.C. performed the simulations, developed the computational pipeline, curated and analyzed the data, and prepared the figures. M.C. and J.F.R.M. interpreted the results, wrote the manuscript, critically revised the work, and approved the final version.

Competing interests. M.C. and J.F.R.M. are inventors on a patent application related to the in silico methodology described in this work. The patent application is owned by Politecnico di Milano. The authors declare no other competing interests.

References

- [1] International Council for Harmonisation. ICH E14: The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs (2005). URL https://database.ich.org/sites/default/files/E14_Guideline.pdf. ICH Harmonised Tripartite Guideline.
- [2] International Council for Harmonisation. ICH S7B: The non-clinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals (2005). ICH Harmonised Tripartite Guideline, Step 5.
- [3] Darpo, B. *et al.* Results from the IQ-CSRC prospective study support replacement of the thorough QT study by QT assessment in the early clinical phase. *Clinical Pharmacology & Therapeutics* **97**, 326–335 (2015).
- [4] Garnett, C. *et al.* Scientific white paper on concentration-QTc modeling. *Journal of Pharmacokinetics and Pharmacodynamics* **45**, 383–397 (2018).
- [5] International Council for Harmonisation. ICH harmonised guideline E8(R1): General considerations for clinical studies (2016).
- [6] Friedman, L. M., Furberg, C. D., DeMets, D. L., Reboussin, D. M. & Granger, C. B. *Fundamentals of Clinical Trials* 5 edn (Springer, 2015).
- [7] Mohs, R. C. & Greig, N. H. Drug development: success rates, timelines, and costs. *Frontiers in Pharmacology* **8**, 355 (2017).
- [8] European Medicines Agency. Guideline on strategies to identify and mitigate risks for first-in-human and early clinical trials with investigational medicinal products (EMA/CHMP/SWP/28367/07 Rev. 2) (2022).
- [9] Polak, S. & Wiśniowska, B. Cardiac safety of oncology drugs: new perspectives for new challenges. *Expert Opinion on Drug Metabolism & Toxicology* **15**, 651–663 (2019).
- [10] European Commission. Directive 2010/63/EU on the protection of animals used for scientific purposes (2010).
- [11] Viceconti, M., Henney, A. & Morley-Fletcher, E. In silico clinical trials: how computer simulation will transform the biomedical industry. *International Journal of Clinical Trials* **3**, 37–46 (2016).
- [12] Rodríguez, B., Passini, E., Britton, O. J. *et al.* Cardiac safety in drug discovery and development: new paradigms for cardiac risk assessment. *Progress in Biophysics and Molecular Biology* **159**, 81–90 (2020).
- [13] Colatsky, T. *et al.* The comprehensive in vitro proarrhythmia assay (CiPA) initiative—update on progress. *Journal of Pharmacological and Toxicological*

Methods **81**, 15–20 (2016).

- [14] Crumb, W. J., Vicente, J., Johannesen, L. & Strauss, D. G. An evaluation of 30 clinical drugs against the comprehensive in vitro proarrhythmia assay (CiPA) proposed ion channel panel. *Journal of Pharmacological and Toxicological Methods* **81**, 251–262 (2016).
- [15] Li, Z. *et al.* Improving the in silico assessment of proarrhythmia risk by combining hERG (IKr) channel–drug binding kinetics and multichannel pharmacology. *Circulation: Arrhythmia and Electrophysiology* **10**, e004628 (2017).
- [16] Dutta, S. *et al.* Optimization of an in silico cardiac cell model for proarrhythmia risk assessment. *Frontiers in Physiology* **8**, 616 (2017).
- [17] Li, Z. *et al.* Assessment of an in silico mechanistic model for proarrhythmia risk prediction under the CiPA initiative. *Clinical Pharmacology & Therapeutics* **105**, 466–475 (2019).
- [18] Amuzescu, B. *et al.* Evolution of mathematical models of cardiomyocyte electrophysiology. *Mathematical Biosciences* **334**, 108567 (2021).
- [19] Sarkar, A. X., Christini, D. J. & Sobie, E. A. Exploiting mathematical models to illuminate electrophysiological variability between individuals. *The Journal of Physiology* **590**, 2555–2567 (2012).
- [20] Britton, O. J. *et al.* Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. *Proceedings of the National Academy of Sciences* **110**, E2098–E2105 (2013).
- [21] Muszkiewicz, A. *et al.* Variability in cardiac electrophysiology: Using experimentally-calibrated populations of models to move beyond the single virtual physiological human paradigm. *Progress in Biophysics and Molecular Biology* **120**, 115–127 (2016).
- [22] Passini, E. *et al.* Human in silico drug trials demonstrate higher accuracy than animal models in predicting clinical pro-arrhythmic cardiotoxicity. *Frontiers in Physiology* **8**, 668 (2017).
- [23] Llopis-Lorente, J. *et al.* Combining pharmacokinetic and electrophysiological models for early prediction of drug-induced arrhythmogenicity. *Computer Methods and Programs in Biomedicine* **242**, 107860 (2023).
- [24] Brennan, T., Fink, M. & Rodriguez, B. Multiscale modelling of drug-induced effects on cardiac electrophysiological activity. *European Journal of Pharmaceutical Sciences* **36**, 62–77 (2009).

- [25] Zenzemi, N. *et al.* Computational assessment of drug-induced effects on the electrocardiogram: from ion channel to body surface potentials. *British Journal of Pharmacology* **168**, 718–733 (2013).
- [26] Bi, X., Zhang, S., Jiang, H. & Wei, Z. A multi-scale computational model for the rat ventricle: Construction, parallelization, and applications. *Computer Methods and Programs in Biomedicine* **208**, 106289 (2021).
- [27] Wiśniowska, B. & Polak, S. Virtual clinical trial toward polytherapy safety assessment: combination of physiologically based pharmacokinetic/pharmacodynamic-based modeling and simulation approach with drug-drug interactions involving terfenadine as an example. *Journal of Pharmaceutical Sciences* **105**, 3415–3424 (2016).
- [28] Costabal, F. S., Matsuno, K., Yao, J., Perdikaris, P. & Kuhl, E. Machine learning in drug development: characterizing the effect of 30 drugs on the QT interval using gaussian process regression, sensitivity analysis, and uncertainty quantification. *Computer Methods in Applied Mechanics and Engineering* **348**, 313–333 (2019).
- [29] Aguado-Sierra, J. *et al.* Virtual clinical QT exposure-response studies—a translational computational approach. *Journal of Pharmacological and Toxicological Methods* **126**, 107498 (2024).
- [30] Dominguez-Gomez, P. *et al.* Fast and accurate prediction of drug induced proarrhythmic risk with sex specific cardiac emulators. *npj Digital Medicine* **7**, 380 (2024).
- [31] Mirams, G. R. *et al.* Prediction of thorough QT study results using action potential simulations based on ion channel screens. *Journal of Pharmacological and Toxicological Methods* **70**, 246–254 (2014).
- [32] International Council for Harmonisation. ICH E14 Questions & Answers (R3): The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs (2015).
- [33] European Medicines Agency. Guideline on the assessment of the QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs (2015). URL https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-assessment-qt-qtc-interval-prolongation-proarrhythmic-potential-non-antiarrhythmic-drugs_en.pdf. EMA/CHMP/ICH/2/04 Rev.2, adopted 25 June 2015.
- [34] ICH. E14/S7B clinical and nonclinical evaluation of QT/QTc interval prolongation and proarrhythmic potential—questions and answers (2020). Step 4 version.

- [35] International Council for Harmonisation. E14/S7B questions and answers: Clinical and nonclinical evaluation of QT/QTc interval prolongation and proarrhythmic potential (2022). ICH Harmonised Guideline, Step 4, adopted 21 February 2022.
- [36] Kramer, J. *et al.* MICE models: superior to the hERG model in predicting Torsade de Pointes. *Scientific Reports* **3**, 2100 (2013).
- [37] Escobar, F. *et al.* Experimentally validated modeling of dynamic drug-hERG channel interactions reproducing the binding mechanisms and its importance in action potential duration. *Computer Methods and Programs in Biomedicine* **254**, 108293 (2024).
- [38] Association of the British Pharmaceutical Industry (ABPI). Guidelines for phase i clinical trials: 2018 edition. Tech. Rep., The Association of the British Pharmaceutical Industry, London, UK (2018).
- [39] U.S. Food and Drug Administration (FDA). Expansion cohorts: Use in first-in-human clinical trials to expedite development of oncology drugs and biologics. guidance for industry. Tech. Rep., U.S. Department of Health and Human Services, Food and Drug Administration (2021).
- [40] DeMarco, K. R., Bekker, S., Clancy, C. E., Noskov, S. Y. & Vorobyov, I. Digging into lipid membrane permeation for cardiac ion channel blocker d-sotalol with all-atom simulations. *Frontiers in Pharmacology* **9**, 26 (2018).
- [41] Elliott, J., Mainardi, L. & Rodriguez Matas, J. F. Cellular heterogeneity and repolarisation across the atria: an in silico study. *Medical and Biological Engineering and Computing* **60**, 3153–3168 (2022).
- [42] McKay, M. D., Beckman, R. J. & Conover, W. J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239–245 (1979).
- [43] Passini, E. *et al.* Drug-induced shortening of the electromechanical window is an effective biomarker for in silico prediction of clinical risk of arrhythmias. *British Journal of Pharmacology* **176**, 3819–3833 (2019).
- [44] O’Hara, T., Virag, L., Varro, A. & Rudy, Y. Simulation of the undiseased human cardiac ventricular action potential: model formulation and experimental validation. *PLoS Computational Biology* **7**, e1002061 (2011).
- [45] Costi, M., Ferrero, J. M. & Matas, J. F. R. Toward standardized methodologies for drug-induced proarrhythmia classification: an in silico proof-of-concept. *Computational and Structural Biotechnology Journal* **35**, 0041 (2026).

- [46] Qu, Y. *et al.* Action potential recording and pro-arrhythmia risk analysis in human ventricular trabeculae. *Frontiers in Physiology* **8**, 1109 (2017).
- [47] Heidenreich, E. A., Ferrero, J. M., Doblaré, M. & Rodriguez, J. F. Adaptive macro finite elements for the numerical solution of monodomain equations in cardiac electrophysiology. *Annals of Biomedical Engineering* **38**, 2331–2345 (2010).
- [48] Myers, J. H., Stirling, M. C., Choy, M., Buda, A. J. & Gallagher, K. P. Direct measurement of inner and outer wall thickening dynamics with epicardial echocardiography. *Circulation* **74**, 164–172 (1986).
- [49] Zhang, M.-K., Zhang, Z., Xue, H., Fan, L. & Wen, Y. Microvascular rarefaction and myocardial fibrosis in hypertrophic obstructive cardiomyopathy: a histopathological comparison of pediatric and adult patients. *The Heart Surgery Forum* **25**, E042–E047 (2022).
- [50] MacCannell, K. A. *et al.* A mathematical model of electrotonic interactions between ventricular myocytes and fibroblasts. *Biophys J* **92**, 4121–4132 (2007).
- [51] Kléber, A. G. & Jin, Q. Coupling between cardiac cells—an important determinant of electrical impulse propagation and arrhythmogenesis. *Biophysics Reviews* **2**, 031301 (2021).
- [52] Celotto, C. *et al.* The frequency of atrial fibrillatory waves is modulated by the spatiotemporal pattern of acetylcholine release: a 3d computational study. *Frontiers in Physiology* **14**, 1189464 (2024).
- [53] 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. *Cardiovascular Research* **58**, 32–45 (2003).
- [54] Geselowitz, D. B. On the theory of the electrocardiogram. *Proceedings of the IEEE* **77**, 857–876 (1989).
- [55] Gima, K. & Rudy, Y. Ionic current basis of electrocardiographic waveforms: a model study. *Circulation Research* **90**, 889–896 (2002).
- [56] Malik, M. & Batchvarov, V. Measurement, interpretation, and clinical potential of QT dispersion. *Journal of the American College of Cardiology* **36**, 1749–1766 (2000).
- [57] Postema, P. G., Jong, J. S. D., der Bilt, I. A. V. *et al.* Accurate electrocardiographic assessment of the QT interval: teach the tangent. *Heart Rhythm* **5**, 1015–1018 (2008).

- [58] Robyns, T. How to: measure the QT interval? *Cardiovascular Genomics Insight* **9**, — (2024).