Validation and Trustworthiness of Multiscale Models of Cardiac Electrophysiology

Pras Pathmanathan* and Richard A. Gray

Office of Science and Engineering Laboratories, Center for Devices and Radiological Health, U.S. Food and Drug Administration, Silver Spring, MD, United States

Computational models of cardiac electrophysiology have a long history in basic science applications and device design and evaluation, but have significant potential for clinical applications in all areas of cardiovascular medicine, including functional imaging and mapping, drug safety evaluation, disease diagnosis, patient selection, and therapy optimisation or personalisation. For all stakeholders to be confident in model-based clinical decisions, cardiac electrophysiological (CEP) models must be demonstrated to be trustworthy and reliable. Credibility, that is, the belief in the predictive capability, of a computational model is primarily established by performing validation, in which model predictions are compared to experimental or clinical data. However, there are numerous challenges to performing validation for highly complex multi-scale physiological models such as CEP models. As a result, credibility of CEP model predictions is usually founded upon a wide range of distinct factors, including various types of validation results, underlying theory, evidence supporting model assumptions, evidence from model calibration, all at a variety of scales from ion channel to cell to organ. Consequently, it is often unclear, or a matter for debate, the extent to which a CEP model can be trusted for a given application. The aim of this article is to clarify potential rationale for the trustworthiness of CEP models by reviewing evidence that has been (or could be) presented to support their credibility. We specifically address the complexity and multi-scale nature of CEP models which makes traditional model evaluation difficult. In addition, we make explicit some of the credibility justification that we believe is implicitly embedded in the CEP modeling literature. Overall, we provide a fresh perspective to CEP model credibility, and build a depiction and categorisation of the wide-ranging body of credibility evidence for CEP models. This paper also represents a step toward the extension of model evaluation methodologies that are currently being developed by the medical device community, to physiological models.

Keywords: credibility, calibration, validation, computational modeling, cardiac models

INTRODUCTION

One of the most remarkable properties of the natural world is that is it can be understood using mathematical equations—a property described by Eugene Wigner as "the unreasonable effectiveness of mathematics in the natural sciences." Once the appropriate mathematical groundwork had been developed, it became possible to describe intricate and multi-faceted physical

phenomena using relatively simple mathematical equations, e.g., fluid flow, deformation of solid bodies, electromagnetic wave propagation, and phenomena at widely different scales from atoms to galaxies. Computational models, which are mathematical models solved by means of a computer, can be used to solve governing equations underlying complex systems and simulate their behavior. The remarkable predictive capability of computational models based on the fundamental laws of physics has enabled such models to be routinely used in a multitude of engineering applications.

Biology, in contrast to physics, is less easily characterized by simple or small numbers of mathematical equations. Primarily, this is due to the complexity and variability in biological processes which makes them inherently non-linear, multi-disciplinary and multi-scale. While computational models of human physiological processes have been developed and refined for decades, they are not as predictive as computational models in engineering, and likely never will be. Nevertheless, biomedical computational models have without doubt the potential for revolutionizing medicine just as physics-based models have forever changed research, design, and evaluation in engineering.

One field which holds considerable promise for clinical applications is cardiac modeling, owing to the maturity of the field (Trayanova, 2011) (Winslow et al., 2012) and the fact that heart disease is the leading cause of the death in the industrialized world. Computational cardiac models can simulate the electrophysiology and/or mechanical deformation of cardiac myocytes, tissue, or the whole heart. This paper is focused on cardiac electrophysiological (CEP) models. **Figure 1** illustrates the typical components to a CEP model, which are usually multi-scale, containing as a fundamental unit a cellular model of myocyte EP activity. Such "cell models," of which over a 100 have been published of varying complexity and for a range of mammalian species, are typically sets of ordinary differential equations (ODEs), and predict the action potential (AP) and many other cellular and sub-cellular quantities. Notable recent human cell models include (Iyer et al., 2004; ten Tusscher et al., 2004; ten Tusscher and Panfilov, 2006; Grandi et al., 2010) and (O'Hara et al., 2011). For reviews of single cell models, see (Fink et al., 2011; Noble, 2011; Noble et al., 2012). Cell models are often composed of multiple sub-models, for different ion channels, pumps and exchangers or representing subcellular processes such as calcium handling. These sub-models are usually also systems of ODEs. To simulate electrical wave propagation, including arrhythmic activity, in tissue or the whole heart, cell models are coupled to partial differential equations (PDEs) known as the "bidomain" or "monodomain" equations (Clayton et al., 2011; Franzone et al., 2014; Lopez-Perez et al., 2015); see **Figure 1**. A further extension is to model the heart embedded in the torso, which allows for simulation of the electrocardiogram (ECG) (Richards et al., 2013; Zemzemi et al., 2013). In recent years various imaging, modeling and computational advances have enabled *patient-specific* heart models to be generated using clinical data (see e.g., Smith et al., 2011; Chabiniok et al., 2016 for discussions). Anatomical personalisation using clinical imaging data is most common (e.g., Arevalo et al., 2016), although personalisation of functional (Chen et al., 2016) or material

(Aguado-Sierra et al., 2011) parameters using clinical data has also been performed. Patient-specific models can be used to make patient-specific clinical predictions and represent an important step forward toward precision medicine.

All types of CEP model—ion channel or subcellular models, cell models, tissue, and organ-level models—have proved to be powerful tools complementing experiment in basic cardiac electrophysiological research (Plank et al., 2009), for understanding mechanisms behind both normal rhythm and cardiac arrhythmias. However, CEP models also have potential applications in all aspects of cardiovascular medicine, including device design and evaluation, functional imaging and mapping, drug safety evaluation, disease diagnosis, patient selection, and therapy optimisation or personalisation. There are numerous reviews covering the current and potential applications of CEP models; (see e.g., Trayanova and Boyle, 2014; Davies et al., 2016; Niederer and Smith, 2016). However, one aspect of the modeling which has been inadequately covered in the current cardiac modeling literature is rationale for when and why cardiac models can be *trusted*.

The **credibility** of a computational model has been defined as the belief in the predictive capability of the model for a given intended use (ASME, 2017) or the willingness of people to trust a model's predictions (Patterson and Whelan, 2017). Typically, credibility of a computational model is founded upon **validation** results. Validation is the process of testing a model by comparing model predictions to experimental or clinical data. (A more precise definition is provided below). However, other types of evidence (non-validation evidence) can also be used to argue that a model is sufficiently credible for its intended use. As we explain below, the credibility of CEP models is typically founded upon a wide range of factors, and consequently it can be very unclear, or a matter for debate, the extent to which a cardiac model can be trusted for a given application. In fact, many papers in the literature leave implicit why, and to what extent, the model predictions can be considered credible.

The aim of this article is to clarify and discuss reasons that could support the trustworthiness of CEP models. We will do so by reviewing different types of evidence that have been presented to support CEP model credibility, or are otherwise relevant to the assessment of credibility, hereafter referred to as CEP model **credibility evidence**. The review will include: (i) discussion of common practice regarding CEP model validation; (ii) examples of strategies taken for performing CEP model validation; and (iii) discussion of other credibility evidence for CEP models, including historical evidence that often *implicitly* supports simulation studies. The review will *not* aim to evaluate specific cardiac models or in any way judge the *quality* of any validation results or other evidence. Such decisions require difficult judgements based on careful consideration of all available evidence, in the context of the precise application that model is to be used for (including assessment of the risks associated with inaccurate predictions NASA, 2009) and are therefore beyond the scope of this review. In other words, we are *not* claiming or implying that any CEP model "is" or "is not" trustworthy; nor do we argue that any modeling approach or process is or is not trustworthy. Instead, our focus will be on *types*

FIGURE 1 | Components of a multiscale cardiac electrophysiology (CEP) model. **(Left)**: equations and sample output for a Hodgkin-Huxley formulation of the rapid sodium current. Multiple such sub-cellular models can be used to define a cell model. **(Center)**: schematic of sub-cellular processes included in a hypothetical cell model, together with the differential equation governing the transmembrane voltage, and sample output. Cell models differ in their formulation of the ionic current I_{ion} and can be made up of dozens of ordinary differential equations. **(Right)**: Cell models can be incorporated into the bidomain equations and solved on a computational mesh of the heart [top right: high-resolution rabbit biventricular mesh of Bishop et al. (2010)], to simulate normal or arrhythmic cardiac activity (bottom right).

of evidence that could, in principle, support the trustworthiness of a model for a given intended use.

In previous work we advocated that engineering model assessment approaches known as verification, validation, and uncertainty quantification (VVUQ) could be important in the advancement of cardiac CEP modeling (Pathmanathan and Gray, 2013) and explored verification (Pathmanathan and Gray, 2014) and uncertainty quantification (UQ) for CEP models (Pathmanathan et al., 2015). This paper continues this line of work by focusing on validation-related activities. Only activities related to comparison of the model to the real world are within the scope of this review. Therefore, activities such as code verification, calculation verification, and sensitivity analysis, while important for overall assessment of credibility and receiving increasing interest in the field (Sobie, 2009; Niederer et al., 2011; Chang et al., 2015) are outside the scope of this paper. Additionally, while uncertainty quantification is related to validation as will be described in section Why Trust a Computational Model?, research on the process of performing uncertainty quantification with CEP models is also outside the scope of the review, though this is also receiving increasing recent interest in CEP; (see e.g., Konukoglu et al., 2011; Johnstone et al., 2016; Chang et al., 2017).

In fact, there is enormous current interest across computational science in methods and best practices for demonstrating or evaluating the reliability of computational models (National Research Council, 2012). The medical device community is collaborating on a Standard for assessing credibility of computational models for medical device applications (ASME, 2017). However, this Standard and related

medical device Guidances (Food and Drug Administration, 2016) are motivated by traditional "physics-based" engineering models in biomedical applications [e.g., models based on solid mechanics (Pelton et al., 2008), fluid dynamics (Stewart et al., 2012), or electromagnetism (Angelone et al., 2010)]. The relevance of such approaches to the evaluation of complex physiological models such as CEP models is limited. In particular, while both (Food and Drug Administration, 2016) and (ASME, 2017) address validation, they do not account for the possibility of multiple sources of different types of validation evidence, or other evidence which could support credibility. In this paper, we demonstrate how a large body of evidence will generally support a CEP model. By exposing and discussing this wide range of potential credibility evidence for CEP models, this paper is a necessary first step toward the extension of the above approaches to cardiac and other physiological models.

The paper is organized as follows. In section Why Trust a Computational Model? we categorize and discuss different types of credibility evidence, and discuss validation in detail. Section Credibility of CEP Models at Different Spatial Scales then reviews credibility evidence for CEP models organized by spatial scale. Section Discussion summarizes and discusses our conclusions.

WHY TRUST A COMPUTATIONAL MODEL?

Figure 2 provides an overview of the concepts discussed throughout this section. Various types of rationale could be used to argue for the credibility of a computational model. The following are three distinct categories of evidence that could support *some* level of confidence in a model:

FIGURE 2 | Illustration of how a multiscale CEP model may be supported by multiple sources of credibility evidence (that is, evidence relevant when assessing the credibility of the model), taken from model evaluation at multiple scales. The overall model (i.e., organ-level model), the underlying cell models (here it is assumed that the organ-level model incorporates two different cell models, one for epicardial tissue, one for endocardial), and all underlying sub-cellular models may all be supported by the different types of evidence presented in the right of the figure. See section Why Trust a Computational Model? for full discussion.

Category 1: Evidence related to the **validity of assumptions underlying the model governing equations**, together with evidence related to the **accuracy/fidelity of model parameter values** and other inputs. These are grouped together because if the equations are considered appropriate, but there is no confidence that the parameters are accurate, then there will be little confidence in model predictions, and vice versa.

Category 2: **Calibration evidence**. Calibration is the process of tuning, fitting or optimizing parameters in a model so that the model results match experimental or clinical data. Calibration is primarily performed to determine model parameter values that cannot be directly measured. However, if the results from calibration demonstrate a good match between model and the experimental/clinical data, the results could potentially *also* be used as evidence for credibility of the model.

Category 3: Evidence generated from testing the predictive capability of the *completed* model. This includes **validation evidence**, that is, comparison of model predictions to independent real-world data *not used* in the construction of the model (Roache, 2009). Validation is discussed in detail later in this section.

These are distinct types of evidence and may provide very different levels of confidence in a model. The first category is based on model equations, assumptions and parameter values, but not on actual model outputs or results, with no actual testing of the model. This category includes historical evidence supporting the governing equations that were used in the model.

It also includes evidence regarding the quality of data used to determine model parameters. The second category is based on comparing model outputs with experimental/clinical data, but allows for model parameters to be altered for the model to match the data. Calibration results are regularly used, either implicitly and explicitly, as evidence for credibility of biological computational models. If a model's parameters can be chosen so that the model reproduces certain experimental data, this ability to fit the data or to reproduce phenomena could be used in support of the model—especially when a complex phenomenon is reproduced using a model with only a few parameters. The last category—validation and related evidence—is obviously the strongest test of the model: it assumes the model has been completely defined and *then* its ability to reproduce the real world is tested.

For many applications—in particular the basic science applications of hypothesis generation and mechanistic insight—use of a model that has no supporting validation evidence may be perfectly appropriate. Mathematical models in biology can be thought of a means in which existing knowledge or hypotheses are integrated (Brodland, 2015), in which case running a simulation is simply an efficient means of determining the logical consequences of those knowledge/hypotheses, impossible through mental deduction alone for complex systems. This is one of the reasons why mathematical modeling has proven a successful complement to experiments in understanding biological processes (Noble, 2011). However, when a model is to be used in decision-making, and in particular for

high-risk applications such as safety-critical clinical applications, validation becomes very important. (Carusi et al., 2012) provide a discussion on the meaning of CEP models as representations. Patterson and Whelan (2017) provide an excellent general discussion on models as representations vs. as predictive tools, and propose a high-level framework for deciding how to evaluate models along this spectrum.

Validation

Validation has been described as "*the assessment of the accuracy of a computational simulation by comparison with experimental data*" (Oberkampf et al., 2004). One definition initially proposed by the Department of Defense (DoD) and commonly used by the engineering community and elsewhere (Oberkampf et al., 2004; Roache, 2009; National Research Council, 2012), including increasingly within the medical devices community (Food and Drug Administration, 2016), (ASME, 2017), is: "*the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.*" This definition emphasizes that the validation process is dependent on the specific "intended use" of the model, also referred to as the "model application" or the "**context of use**" **(COU)**. (Roache, 2009) provide a good discussion of the DoD definition of validation, and explains how many different interpretations of it are made even within the engineering community. Moreover, there is no inter-disciplinary consensus on a precise definition of validation, and different communities may have very different understandings of what constitutes model validation. (Bellocchi et al., 2010) list 20 definitions of model validation proposed between 1960 and 2010. (Eddy et al., 2012) describe a categorisation of validation used in the health economics and outcomes community. (Patterson and Whelan, 2017) describe a broad concept of validation of biological models, which includes but is more expansive than the engineering/DoD understanding of validation. (Viceconti, 2011) refer to model "falsification," rather than validation, based on the contention that models can only be invalidated (falsified). One common feature of most of the different interpretations of validation is that validation must involve new data not used in the construction of the model, i.e., "calibration is not validation" (Roache, 2009).

The DoD definition arguably presents a contradiction regarding validation of CEP models—especially *cellular* models—because they are typically developed as **general-purpose models**, i.e., without a specific COU in mind. When novel cell models are published, predictions of model outputs and derived quantities—for example, action potential shape, action potential duration (APD) restitution, ionic concentration transients and others—are usually compared to experimental data. This arguably does not constitute validation according to the DoD definition since no COU (intended use) has been prescribed. Incidentally, this could be considered a limitation of the DoD definition; see (Roache, 2009) for a discussion. Note though that regardless of the definition of validation, it is vital to recognize the importance of the COU in evaluation of a computational model. In particular, the COU must be specified for the "final" evaluation of a model, and any claim that a general-purpose model is a "validated model" cannot

be justified, since it is the COU that determines what level of agreement between model and experiment can be deemed acceptable (Roache, 2009; National Research Council, 2012).

To distinguish between different types of evidence, we introduce the terminology **general validation evidence** to describe scientific evidence obtained by comparing model predictions with real world data *when no particular COU has been specified*. This could also have been called "baseline validation." Examples of general validation for CEP models include initial validation of a novel general-purpose cell model (discussed in detail in section Cell Models), general testing vs. experimental/clinical data of previously published cell models (regularly carried out in the cardiac modeling community), and the comparison against data of activation patterns predicted by general-purpose ventricular, atrial or whole-heart models (discussed in section Organ-Level Models). There is almost unlimited scope for such evaluation, since modern CEP cell models are very complex, and therefore there is an ever-growing volume of literature incorporating general validation of CEP models.

We define **COU-driven validation evidence** using the DoD definition, as scientific evidence obtained by comparing model predictions with real world data for the purposes of evaluating the predictive capability of the model for a specific, prescribed, application (COU) of the model. A simple example of this would be comparing APDs of a model to experimental values, when the COU is prediction of drug effects on APD. Another example is comparison of whole-heart model predictions of number of phase singularities during ventricular fibrillation, against clinical data, when the COU is to use the model to understand mechanisms behind ventricular fibrillation (see section Organ-Level Models). We include in this category validation of model-derived quantities, including: drug pro-arrhythmic risk indices (see sections Cell Models and Organ-Level Models); sudden cardiac death (SCD) risk indices (see section Organ-Level Models); and ablation targets (see section Organ-Level Models).

Note that a terminology complication can arise when considering validation of patient-specific models, which are often generated using a workflow that may be mostly or fully automated. One could distinguish between validation of the simulation software *only*, and validation of model predictions using the full *workflow*; there is therefore a potential for different interpretations of what constitutes "model validation" in this context. In this paper, we will include validation of the full workflow (for example, evaluation of the predictive ability of a workflow that takes in patient imaging data and outputs a clinical prediction) within our broad interpretation of model validation.

Comparator, Quantities of Interest, and Method of Comparison

Validation involves comparison of model predictions with real world data of some form. (Note that comparison against results of a *different* computational model is generally not considered validation, but see Roache, 2009 for a discussion). The **comparator** is defined as the source of the real world data. For CEP models this is usually experimental or clinical data.

Important aspects of experimental comparators in CEP model validation include species, experimental conditions including temperature, and whether the data is historical (taken from the literature) or obtained from new experiments performed for the purpose of model validation. Important aspects of clinical comparators include patient demographics, patient cardiac myopathies, and co-morbidities. For patient specific CEP models that make patient specific clinical predictions, the validation comparator has to be clinical data taken from the same patient (distinct to the data used for personalisation of the model). Regardless of whether the comparator is experimental or clinical data, there are often significant challenges to obtaining high quality data, especially *in vivo* data under physiological conditions, which can impose severe constraints on the ability to perform high quality validation. These experimental/clinical challenges are covered elsewhere in the CEP literature, and therefore will not be a main focus of the present review.

Another important aspect of validation is which outputs of the model, or derived quantities—here referred to as **quantities of interest (QOIs)**—are compared to the real-world data. Commonly validated QOIs for cell models include transmembrane voltage and the APD restitution curve. For whole-heart models, validation QOIs can be global (e.g., the ECG) or local (e.g., activation patterns). Validation using global QOIs only provides indirect evidence on the credibility of local QOIs.

There are various possibilities for the **method of comparison** between the model and comparator. (Oberkampf et al., 2004) provide a good introduction to this topic; here we only provide a very brief overview. The comparison can be qualitative (often the case in physiological modeling) or quantitative. If quantitative, the comparison could take into account experimental error, model uncertainty, both, or neither. Model uncertainty is accounted for by performing uncertainty quantification (UQ), where uncertainty in model parameters (due to, for example, measurement uncertainty or inherent physiological variability) is quantified using probability distributions, and then the resultant uncertainty in the QOI(s) are computed (Smith, 2013; Mirams et al., 2016). Various validation metrics for quantifying the difference between experimental data and model predictions taking into account error estimates and simulation uncertainty have been proposed in the engineering literature; (see e.g., Oberkampf and Barone, 2006). For some CEP model-derived outputs such as risk indices or model-based biomarkers, other analytic or statistical comparison methods (different to those used in traditional model validation) may be appropriate, such as measures of specificity and specificity, receiver operating characteristic (ROC) curves, biomarker validation methods, etc.

Sometimes a CEP model is stated as matching known physiological phenomena, for example in statements such as "the AP shows the characteristic spike notch dome architecture found for epicardial cells" (ten Tusscher et al., 2004) or discussion of re-entrant waves breaking up into sustained fibrillation under pro-arrhythmic conditions (Krishnamoorthi et al., 2014). This is perhaps not validation *per se*, as there is no explicit comparator— or more precisely, arguably not validation according to the engineering/DoD understanding of validation; it is arguably

"epistemic validation" using the broader definition of Patterson and Whelan (2017). Nevertheless, it is important and relevant evidence for assessing the model's predictive ability for a COU. This type of evidence, which we will refer to as **reproduced phenomena**, may be especially important in evaluation of biological models since biological systems exhibit emergent phenomena, and therefore a powerful test of a model is whether such it predicts such phenomena.

Validation of Multiscale Models

For multiscale models we can distinguish between evidence at different spatial scales, and in particular at which scales validation was performed (see **Figure 2**). For a multiscale model of the whole-heart, there may be validation evidence available for model sub-components (i.e., all sub-cellular models and the cell model), and/or for the system-as-a-whole (whole-heart model). If validation is only performed for sub-models but *not* the overall system, credibility of system-level predictions is founded (perhaps implicitly) on the sub-model validation results *and* belief in the theory underlying how sub-models interact. For example, most cardiac cell models assume that ionic currents are independent and can therefore simply be added together. System-level validation may be especially important with physiological models, since physiological systems exhibit emergent behavior that cannot be predicted from understanding all sub-system behavior. "Hierarchical validation," in which validation is performed for all model sub-components, sub-systems *and* the entire system, is recommended in the engineering validation literature so that the model provides the "right answer for the right reasons" (Hills et al., 2008).

Often, validation is performed at one scale to provide confidence that the model is sufficiently credible for it to be used as a sub-model in a larger scale (e.g., develop a cell model, perform validation of cell model, and then proceed to tissue model if validation results are favorable). Even if this is the case, the sub-model validation results may be relevant in evaluation of the final model for a COU.

It should now be clear how a CEP model may be supported by **multiple sources of credibility evidence, taken from model evaluation at multiple scales** (see **Figure 2**). **Table 1** lists different sources of evidence and provides examples for ion channel, cell and organ-level models. We reiterate that we are not making any assertions regarding what evidence is *necessary* when assessing cardiac models for a COU. Our motivation is simply to describe how multiple sources of evidence may exist and be relevant when assessing the credibility of a CEP model for a specific COU. Confidence in a model tends to increase with the body of evidence available to support it (Patterson and Whelan, 2017). Therefore, when a complex model is evaluated, ideally the model should be treated as a "glass box" (the opposite of a "black box"), so that the most informed decision is made. Any or all of the types of evidence in **Table 1** may be relevant in glass box cardiac model evaluation. The most important source of evidence for a whole-heart model will likely be organ-level COU-driven validation evidence, if available. Strong validation results of the full model, if highly "applicable" (Pathmanathan et al., 2017) to the COU, reduce the relative importance of the other factors

TABLE 1 | Different types of evidence relevant to the credibility of a cardiac EP model, with ion channel, cell, and organ-level examples.

Category	Type of credibility evidence	Examples		
		Ion channel	Cell model	Organ-level model
Category 1	Evidence regarding validity of model assumptions or supporting the model formulation	Successes of Hodgkin-Huxley formulation for modeling ion channels—see section Ion channel models	Evidence supporting the formulation of cell membrane as a parallel resistor-capacitor electric circuit	The successes of the bidomain equations, in particular predictions made that were later experimentally observed—see section Organ-level models
	Evidence regarding accuracy/fidelity of model parameters/inputs	Evidence supporting accuracy of steady-state inactivation parameters—see section Ion Channel Models	Rationale behind standard choice of membrane capacitance equal to 1 uF/cm^2.	Evidence on fidelity of geometry used and on fidelity of fiber/sheet specification—discussed in section Organ-Level Models.
Category 2	Calibration results	Results showing agreement between ion channel model and experimentally recorded current-voltage relationship when ion channel parameters are calibrated using this data	Results showing agreement between the model action potential and experimental recordings when maximal conductances are tuned to achieve the match	Results showing activation patterns match experiment if fast sodium current maximal conductance (which controls conduction velocity) chosen to maximize agreement
Category 3	Reproduced (emergent) phenomena	Simulation results demonstrating that a rapid sodium current model can exhibit damped oscillations	Simulation results demonstrating that a cell model reproduces action potential spike and dome morphology	Simulation results demonstrating that ECG predicted by a heart and torso model exhibits realistic-looking QRS complex and T wave
	General validation results	Comparison of a general-purpose ion channel model predictions to new voltage-clamp data not used in the construction of the model.	Comparisons of model results with experimental data for a novel general-purpose cell model, e.g., all such results in O'Hara et al. (2011). Discussed in detail in section Cell Models	Excitation patterns of general purpose bi-ventricular model compared to experimental/clinical data. ECG of general-purpose heart and torso model compared to experimental/clinical data.
	COU-driven validation results	Evaluation of a hERG model to predict pharmaceutical pro-arrhythmic risk	Evaluation of a cell model-based biomarker to predict pharmaceutical pro-arrhythmic risk (e.g., CiPA, discussed in section Cell Models)	Number of phase singularities during ventricular fibrillation (VF) compared to clinical data, when the model will be used to understand mechanisms behind VF—see section Organ-Level Models. Clinical evaluation of a whole-heart model which uses patient-specific information to predict optimal ablation targets to terminate arrhythmias—see section Organ-Level Models

(including reducing the need for evidence supporting model assumptions (Patterson and Whelan, 2017).

CREDIBILITY OF CEP MODELS AT DIFFERENT SPATIAL SCALES

We now discuss credibility evidence of CEP models at each of the spatial scales. The scope of the following review is limited to the most common types of CEP model: zero-dimensional models (i.e., systems of ODEs) of ionic channels and of the cell, and tissue/organ models that utilize the monodomain or bidomain formulation. Therefore, models that explicitly represent the spatial structure of ion channels or cardiac myocytes are out of scope of the review, including molecular dynamics models. Due to space limitations, we will only discuss ion channel models; other types of sub-cellular model such as calcium handling models are not included. We only consider models which are at least partially motivated by bio-physical understanding, excluding phenomenological models, or statistical models such as those developed using neural networks or machine learning techniques. We re-iterate that this paper is focused on electrophysiology only; models of cardiac mechanics or hemodynamics are out of scope, although similar principles are expected to apply. Note that the scope of the following review is still quite broad and it is therefore not possible to describe or cite all publications that have performed validation of CEP models. The papers cited below were chosen to provide selected examples of approaches to CEP model validation.

Ion Channel Models

There is a long history of modeling the dynamics of transmembrane ion channels using the Hodgkin-Huxley (HH) formulation (Hodgkin and Huxley, 1952). In the HH formulation, transmembrane current is taken to be the product of

a maximum conductance, dynamic gating variables representing probabilities of channels gates being in an open state, and a driving force. Gating variable dynamics are modeled using ODEs, with dynamics determined by the voltage-dependent steady-state activation/inactivation and voltage-dependent "time constant" relationships for each gating variable (see **Figure 1**, which includes the equations for a HH formulation of the rapid sodium current with three gating variables m, h, and j). The HH formulation has in fact become so integral to cardiac electrophysiology that experimentalists regularly present data by publishing HH-based model parameters. Markov models of ionic currents are a more general formulation. For more details (see e.g., Fink et al., 2011).

Generally speaking, validation of novel ion channel models is *not* common practice (Fink et al., 2011). Here, we are referring to validation of the novel channel in isolation, not as part of a larger cell model. While voltage clamp data is used to develop and calibrate the models, those calibrated models are typically not then tested to new data. In fact, generally simulations are not even performed to show that the models predict the voltage-clamp results that they were based on, and surprisingly, simulations of voltage clamp protocols from which parameters are derived do not necessarily match the original data (Carro et al., 2017). (This can happen for a variety of reasons, such as the assumption of inactivation being much faster than activation not holding). Such observations demonstrate the value of ion channel model evaluation including validation. It can be difficult to determine in publications if results presented correspond to validation, because calibration and validation are often not clearly separated in presentation of results. An example of genuine validation is (Yang et al., 2015), in which validation of a new model of the late sodium current I_{NaL} is performed by comparing model predictions of the I_{NaL} current-voltage relationship under a slow depolarising voltage ramp, against experimental recordings under the same protocol. Another is the L-type calcium current model in O'Hara et al. (2011). As shown in **Figure 3**, validation of the calibrated I_{CaL} model was performed by comparing model with experimental data using an action potential clamp protocol. (Beattie et al., 2017) proposes a novel approach to developing cell-specific models of the rapid delayed rectifier potassium current I_{Kr}. Eight seconds of data using a novel sinusoidal voltage clamp protocol was used to calibrate the cell-specific I_{Kr} models, which were then validated against 5 min of data taken from the same cell, covering a range of voltage clamp protocols.

When no validation evidence is presented, the credibility of novel ion channel models is essentially founded—often implicitly—on a range of other factors, including the maturity of the HH formulation and related historical evidence, calibration evidence, and evidence regarding the accuracy of identified parameters. We discuss these in the remainder of this section.

The model of squid giant axon excitability proposed by Hodgkin and Huxley (Hodgkin and Huxley, 1952) is considered one of the greatest successes in twentieth century biophysics (Häusser, 2000; Schwiening, 2012). This is due to the ability of the relatively simple set of equations to reproduce a variety of phenomena (Häusser, 2000) and the fact that the HH

modeling approach was then successfully applied to a wide range of excitable cells, including cardiac cells (Noble, 1962). The ideas and equations behind the HH model are now standard building blocks in electrophysiology (Schwiening, 2012). HH-based cardiac models have contributed greatly to understanding of cardiac electrophysiology, with various predictions made using cardiac models that were later experimentally verified. Examples include the existence of non-sodium inward currents and stoichiometry of the Na^+/Ca^{2+} exchanger; see (Noble, 2011) for a detailed review. However, despite these successes, there are several caveats that should be stated regarding use of a general HH formulation for a given ion channel. First, for some ion channels and some applications, such as the rapid delayed rectifier potassium current I_{Kr} and drug-binding applications, a Markov model based approach may be more appropriate (Clancy and Rudy, 1999). Additionally, there is still a lack of consensus and ongoing research into a variety of details of specific formulations. For example, for the fast sodium current I_{Na}, while the originally-proposed (Hodgkin and Huxley, 1952) and commonly-used m^3 formulation of activation can be argued to have a justification at the molecular level (Armstrong, 2006), it is unclear how to simultaneously represent the various modes of I_{Na} inactivation (fast, slow, and persistent; Nesterenko et al., 2011). Similarly, for the L-type calcium channel $I_{CaL,}$ there is not a unique approach to simultaneously quantifying both voltage- and calcium-dependent inactivation (Grandi et al., 2010).

Once a HH-based model formulation is proposed and justified, model parameters need to be estimated. This includes parameters representing the voltage-dependent steady-state activation/inactivation and time-constant functions for each gate, which are estimated using voltage clamp data. Evidence on the accuracy of these parameters is important in evaluating model credibility, especially if no validation is performed. However, before asking about the accuracy of parameter values, one can ask if ion channel model parameters are uniquely identifiable from experimental data in the first place. A parameter cannot be claimed to be accurate if it is provably unidentifiable given the data. Although the methodology for nonlinear model identifiability has been extensively studied (Rothenberg, 1971; Jacquez and Greif, 1985; Walter and Pronzato, 1996), their utilization in the field of CEP modeling has been limited. The conditions under which model parameters can be identified has been studied in the context of single current sub-models (Beaumont et al., 1993; Wang and Beaumont, 2004; Lee et al., 2006; Csercsik et al., 2012; Raba et al., 2013) and more recently incorporated into a multi-scale framework using a simplified action potential model (Shotwell and Gray, 2016).

Returning to parameter estimation, voltage-dependant *steady-state* (in)activation relationships for many currents are typically well-approximated using sigmoidal functions using standard voltage clamp protocols, however obtaining data for accurate characterisation of voltage-dependent time constants is considerably more difficult. Assumptions underlying voltage clamp protocols should be well understood by model developers and may be questionable for protocols used to identify certain parameters (e.g., the assumption that inactivation is much faster than activation for protocols used to identify I_{Na} steady state

FIGURE 3 | Calibration and validation of the L-type Ca^{2+} current of O'Hara et al. (2011). Left figures show calibration results (circles/squares/diamonds—experiment; solid lines—simulation), including fitting of steady-state activation and inactivation (top row) and time constants (second row). Right figures are qualitative validation of the formulated I$_{CaL}$ model by comparison of simulation and experiment under an identical action potential clamp. Quantitative validation of peak current is also provided in original paper. (Adapted from Figure 1 of O'Hara et al. (2011) with permission under Creative Commons license).

gating activation parameters Csercsik et al., 2012). Manual recording from single cells using well-established voltage clamp protocols remains the gold standard for obtaining high-quality current-voltage relationships (Elkins et al., 2013). Nevertheless, there are numerous (often "hidden") details regarding the specific procedures and protocols in the laboratory to isolate individual currents and to minimize recording artifacts (e.g., accounting for liquid junction potentials and capacitive compensation). Technical advances has improved the ability to measure rapid transients, however, it is still not possible to characterize activation for the fast sodium current steady-state activation from adult myocytes under physiological conditions (Berecki et al., 2010). Experimental reproducibility and variability between cells also present challenges (Pathmanathan et al., 2015). Regarding the voltage dependence of time constants, there is not even consensus on the appropriate functional forms, unlike steady state parameters. In general, fits to time constant data are much poorer than to steady-state (in)activation data [compare steady state and time constant fits in (e.g., ten Tusscher et al., 2004) or (O'Hara et al., 2011); also see **Figure 3**]. The quality of such fits is rarely quantified.

Overall, if a novel ion channel model is developed but no validation is performed, given the numerous issues described above there may be significant uncertainty regarding the true predictive capability of the ion channel model. This is especially true for simulations using conditions that are quite different to the conditions used for model calibration. Often, however, novel ion channel models are developed as one component of a cell model, and validation is instead performed at that level, as described in the next section.

Cell Models

Regarding validation of cardiac cell models, we first make two remarks. First, it should be noted that the majority of cell models are developed as general-purpose tools, as opposed to for a specific COU. Any initial testing against experimental data of a general purpose cell model therefore falls under the category of "general validation." Second, it can be especially difficult to determine in publications whether results presented are obtained by calibration or are genuinely validation evidence. For example, simulated and experimental action potentials may be presented in papers to demonstrate a close match between simulation and experiment, but it can be unclear if any parameters (e.g., ion channel maximal conductances) were tuned to obtain the match.

There is an enormous range in the extent of general validation performed when novel cell models are published.

They can vary in terms of which model outputs are compared to experiment, which pacing protocols are applied, the source of the experimental/clinical data, and the type of comparison between model and experiment (e.g., qualitative vs. quantitative). As an example, for validation of their human cell model, (ten Tusscher et al., 2004) first present action potential and calcium transient time courses under 1 Hz pacing, stating how the AP reproduces the characteristic spike notch dome of epicardial cells and the calcium transient reproduces the experimentally observed rounded-off triangular shape ("reproduced phenomena" evidence as discussed in section Why Trust a Computational Model?). They then semi-quantitatively compare several AP properties and diastolic/systolic calcium concentration with experiment, qualitatively compare APD restitution and conduction velocity restitution results with experiment, as well as present several other validation-related results, including at the tissue level (after coupling the cell model with the monodomain equations). This is arguably more extensive validation than presented for *most* other cell models. The most comprehensive set of validation tests for a new cell model is, as far as we aware, that presented in the original O'Hara-Rudy-dynamic (ORd) model paper O'Hara et al. (2011), in which validation was performed for all of the following QOIs: AP shape under multiple pacing rates, resting voltage, maximum voltage, maximum upstroke velocity, APD restitution properties (steady state; dynamic; with and without channel-specific blockers; single cell and in tissue), APD alternans and accommodation, AP shape with induced early after-depolarisation (EAD), peak intracellular sodium and calcium ion concentrations at multiple rates, calcium transient at multiple rates, and various current voltage relationships under various voltage/potassium/sodium/calcium clamps. This extensive validation, together with the use of human data for model development, are reasons why the ORd model is one of the most highly regarded of modern cell models, although we emphasize that even this model should not be considered a "validated cell model," both because of the issues with such terminology (see section Validation and Roache, 2009), and also because of certain ways it does not match clinical observations (Mann et al., 2016; Dutta et al., 2017).

In general it is important to note that modern cell models may simulate dozens of quantities (i.e., have dozens of state variables), of which usually only a handful have been directly compared to experimental data; this is certainly true even of the ORd model. Credibility in QOIs not compared to experiment is therefore based on "indirect" validation. It should also be appreciated that most cell models are typically *not* validated using data directly related to the initiation and maintenance of arrhythmias, although there are notable exceptions such as the validation involving EADs in O'Hara et al. (2011) or Nordin and Ming (1995) and involving reentrant waves in ten Tusscher et al. (2004).

The above are all examples of general validation; next we consider validation of cell models for a prescribed COU, i.e., COU-driven validation. For single cell cardiac models, the application (i.e., COU) with the greatest current research interest is prediction of proarrhythmic risk of novel pharmaceutical compounds (Davies et al., 2016). (Davies et al., 2012) develop

an ensemble of 19 cell models calibrated to data from 19 dogs for this COU. For validation, they first compare model predictions of drug effect on action potential shape against experimental data (using various compounds). They then test the ability of the model ensemble to predict—blinded—whether a drug will cause AP shortening, prolongation, or have no effect, on a test set of 53 compounds and using measures of sensitivity, specificity and predictivity. Other CEP model-based biomarkers have also been recently proposed (Mirams et al., 2011; Passini et al., 2017), and have been evaluated against risk classifications scores using test sets of compounds. This application area has matured rapidly, and recently regulators, academia and the pharmaceutical industry have come together in the Cardiac *in vitro* Proarrhythmia Assay (CiPA) program (Cavero and Holzgrefe, 2014; Colatsky et al., 2016). The aim of the CiPA program is to develop a novel framework for assessing proarrhythmic risk. The proposed framework involves a series of predominantly nonclinical assays, one of which utilizes a cardiac cell model to integrate drug ion channel effects to the action potential level. The ORd model is being modified for this purpose, and the ultimate aim is to develop a model-based metric that converts drug ion channel effects into a predictive risk index (Dutta et al., 2017). Twelve drugs with well-characterized risk are being used for model and metric development, and the final metric will be validated (in a blinded fashion) using 16 different drugs with well-characterized proarrhythmic risk.

When no validation evidence is available for a cell model, which may be the case for a novel—or considerably modified—cell model, credibility of model predictions is essentially founded, perhaps implicitly, upon multiple factors. This includes the consensus view that the cell membrane can be modeled as a parallel resistor-capacitor electric circuit (Cole and Moore, 1960; Mauro et al., 1970), together with any evidence supporting credibility of each of the sub-cellular models incorporated (i.e., as discussed in section Ion Channel Models), and any calibration evidence (e.g., ability to reproduce AP shape or characteristics when model parameters are selected accordingly). In this case a lot of subject matter expertise may be required to interpret and to judge reliability of predictions.

Organ-Level Models

Tissue- and organ-level simulations have been used for many years and with great success in basic science applications (Trayanova et al., 2006; Bishop et al., 2009). These models involve the solution of the bidomain or monodomain equations (Tung, 1978; Henriquez, 1993; Neu and Krassowska, 1993; Bourgault et al., 2009), incorporating one or more specific cell models, on a computational mesh that approximates the geometry of interest (which can be a 2D monolayer (ten Tusscher et al., 2004), 3D slab of tissue, the atria (Seemann et al., 2006; Zhao et al., 2013), the ventricles (Plank et al., 2009) or the whole heart (Deng et al., 2012; Baillargeon et al., 2014). It is also possible to model the heart in a conductive medium, such as saline bath or the torso, which allows the electrocardiogram and defibrillation to be simulated (Aguel et al., 1999; Richards et al., 2013; Zemzemi et al., 2013; Okada et al., 2015). Tissue-level parameters that need to be prescribed include intra- and extra-cellular conductivities

(dependent on the local fiber and sheet directions Legrice et al., 1995). For more details (see, e.g., Lopez-Perez et al., 2015), which reviews 60 3D cardiac models developed over the past fifty years.

In fact, the methodology for tissue- and organ-level simulation studies is so well-established that simulation studies are routinely published in which a model is used for EP research but no validation results are presented, and no other rationale for credibility is explicitly presented. For such studies, the credibility of the model predictions is essentially based—implicitly—on the following: (i) confidence in the model governing equations (including historical evidence supporting the bidomain formulation); (ii) confidence in the cell model used; and (iii) the accuracy/fidelity of model parameters and geometrical inputs. We discuss each of these below.

First, we note that bidomain equations have a strong biophysical basis, being mathematically derived through a formal homogenisation of an underlying set of governing equations derived from Maxwell's equations (Neu and Krassowska, 1993). The underlying anatomical and physiological assumptions are mostly considered reasonable, although there remains ongoing research into alternative formulations that may better represent electrical propagation through myocardium, for example the fractional diffusion model of Bueno-Orovio et al. (2014), the alternative homogenisation derived by Richardson and Chapman (2011), or the hyperbolic bidomain model of Rossi and Griffith (2017). The bidomain equations reduce to the monodomain equations under the assumption that the intracellular and extracellular conductivity tensors are aligned. While this is known to not be the case in cardiac tissue, in the absence of extracellular stimuli (such as defibrillatory shocks) solutions of the monodomain equations can be very similar to those of the bidomain (Potse et al., 2006; Clayton et al., 2011). Perhaps the strongest evidence supporting the use of the bidomain equations are the numerous historical examples of quantitative predictions from bidomain simulations that have been reproduced experimentally, including complex phenomena that were predicted by simulation studies and only later observed experimentally. The most famous example regards specific virtual electrode patterns: simulations preceded experiment in predicting that unipolar excitation can result in a "dog-bone" shaped virtual cathode with regions of hyperpolarisation (virtual anode) in the vicinity of the virtual cathode (Sepulveda et al., 1989; Wikswo et al., 1991, 1995). This unexpected phenomenon is the result of the unequal anisotropy ratios of the intracellular and extracellular conductivity tensors. (Wikswo and Roth, 2009) provide a detailed review and numerous other examples of bidomain simulations matching experiment.

Credibility of tissue-level predictions is also dependent on the specific cell model used in the simulations. Credibility of cell models was discussed in section Cell Models. However, note that validation at the cell level does not necessarily imply that simulations will reproduce tissue-level phenomena. For example, (Gray et al., 2013) measured the action potential upstroke shape during propagation and found that it differed from that predicted in tissue simulations using a variety of cell models. (Uzelac et al., 2017) show that current cell models when incorporated into tissue level models do not reproduce the voltage and calcium dynamics of alternans. In addition, it is fairly common to adjust cell model parameters in tissue simulations (e.g., to shorten APD when simulating fibrillation Bishop and Plank, 2012), without any "re-validation" of the modified cell model; for such cases, it is unclear how much the previous cell model validation results can be relied upon. It is also increasingly common to re-calibrate cell model parameters in an organ-level model using data taken from intact tissue, including clinical data (e.g., Keldermann et al., 2008); again, it is unclear the extent that the body of previous validation results holds. We will return to this subject in the discussion.

The third factor especially relevant to model credibility when no validation results are available is the accuracy/fidelity of model parameters and other inputs. In regards to parameters we refer to Clayton et al. (2011), which provides a review of the challenges of estimating parameters in the bidomain equations. Note though that when estimating personalized parameters from clinical data for patient specific models, questions can be raised on the identifiability and accuracy those parameters; see (Chabiniok et al., 2016) for a general discussion. Here we focus on geometrical inputs. In organ-level simulations, an important factor that may require consideration when evaluating credibility is the anatomical fidelity of the computational mesh. There are a range of possibilities, from use of simple truncated ellipsoids (Vetter and McCulloch, 1998) to image-based meshes. Meshes vary in terms of the anatomical detail included. For example, they may include ventricular endocardial structures such as papillary muscles and trabeculae (Bishop et al., 2010); atrial endocardial structures such as fossa ovalis (Seemann et al., 2006); myocardial blood vessels (Bishop et al., 2010); and/or the Purkinje system (Romero et al., 2010; Bordas et al., 2011). The appropriate level of detail for specific applications is not yet clear; in particular there is ongoing research into the role of microstructure on the initiation, maintenance and termination of fibrillation (Bishop and Plank, 2012; Connolly et al., 2017). As well as geometry, there is a question on the fidelity of the prescribed fiber and sheet orientations. This can be estimated using DT-MRI imaging data (Mekkaoui et al., 2012); however DT-MRI data can be noisy due to partial volume effects and sensitive to motion artifacts (Bishop et al., 2009; Dierckx et al., 2009). An alternative approach is to use a "rule-based" method (see **Figure 5**, later, for an example), in which a mathematical algorithm is used to generate fiber and sheet architectures [see e.g., (Potse et al., 2006; Bishop et al., 2010; Bayer et al., 2012) for ventricles or (Krueger et al., 2011; McDowell et al., 2012) for atria], and has been shown to provide results that are very similar to those based on DT-MRI (Bishop et al., 2009; Bayer et al., 2012), but may not correctly capture fine-scale details such as fiber direction near the apex, around vessels or near infarcts. Therefore, either way, there may be considerable uncertainty about the true fidelity of the prescribed fiber/sheet directions, which may impact credibility of predictions of quantities expected to be sensitive to anisotropy.

Next, we move on to validation of organ-level models. The ability to perform validation of such models is of

course heavily constrained by difficulties in obtaining the necessary experimental or clinical data for model validation, and therefore the vast majority of validation of organ-level models has been limited to heart surface potentials. Heart surface potential data can be obtained from a variety of measurement modalities, including transmembrane voltage recorded from glass microelectrodes or using fluorescent dyes (e.g., optical mapping), or extracellular electrograms using electrode plaques, socks, baskets, or other mapping systems (contact and non-contact). These measurements vary in their spatial resolution from a single site to hundreds or thousands of sites. Each modality has its advantages and disadvantages; for example optical mapping provides very high spatial resolution but low voltage fidelity, and is always *ex vivo* for human tissue and only *in vivo* with great difficulty for animal experiments (Dillon et al., 1998). In contrast, extracellular electrograms can be used to obtain *in vivo* data but at lower spatial resolution. With the exception of the transmembrane microelectrodes, all modalities do not directly measure transmembrane voltage, which can lead to difficulties in achieving a like-for-like comparison between simulation and experiment. This can be remedied in the computational model. For example, fluorescent signals from optical mapping are different than transmembrane signals in that they have a longer upstroke (Gray, 1999), which was determined to be a result of photon scattering (Hyatt et al., 2003), which led to the development of CEP models that also simulated fluorescence with scattering to enable like-for-like comparisons (Bishop et al., 2007; Roth and Pertsov, 2009).

Many groups have performed validation of organ-level CEP models using data obtained from these modalities. Here we will provide a few representative examples, to give a flavor of the possibilities for validation of surface potentials or derived quantities. (Relan et al., 2011) describe a framework for the functional personalisation of a porcine biventricular model using *ex vivo* porcine optical mapping data. As shown in **Figure 4A**, following calibration using optical recordings under one pacing protocol, they quantitatively validated predictions of epicardial APD and activation time, using optical recordings from the same heart under various different pacing scenarios. (Rodriguez et al., 2005) investigated the role of structural differences between right and left ventricles in vulnerability to electric shocks in the rabbit heart. The study used a combination of biventicular bidomain simulations and optical recordings from an experimental Langendorff-perfused rabbit heart. The setup enabled various QOIs to be qualitatively compared between simulation and experiment (to support the credibility of simulation-based results of the study), including post-shock transmembrane potential distributions on the epicardial surface, and the probability of tachyarrhythmia induction as a function of shock strength and coupling interval. (Muzikant and Henriquez, 1998) and (Muzikant et al., 2002) compare bidomain predictions with experimental results from the paced *in vivo* canine heart measured using a 528 channel electrode plaque. This study is notable because of the quantitative approach to the validation of spatial patterns, analyzing the root mean squared (RMS) error and Pearson's correlation

FIGURE 4 | Examples of quantitative validation of organ-level models. **(A)** Error maps (i.e., difference between model and experiment; here optical mapping-derived experimental data) for depolarisation time (DT) and APD (top row—pacing on left ventricle epicardium; bottom row—pacing on right ventricular endocardium) (Reproduced with permission from Relan et al., 2011). **(B)** Experimentally measured extracellular potential in mV using electrode plaque (top) compared to predictions of extracellular potential from bidomain simulations (bottom), with difference quantified using Pearson's correlation coefficient (r) and root mean squared (RMS) error. (Reproduced with permission from Muzikant et al., 2002).

coefficient between simulation and experiment, for extracellular potential and conduction velocity; see **Figure 4B**. (Niederer et al., 2010) use patient-specific biventricular electro*mechanical* models to investigate the relationship between the Frank-Starling mechanism and cardiac resynchronisation therapy (CRT) efficacy. To calibrate and validate the electrophysiological part of the electromechanical model, they use patient-specific clinical endocardial data obtaining using the EnSite™ cardiac mapping system. Clinical activation maps during sinus rhythm were used for model calibration, and activation maps under left ventricular pacing were used for validation of the calibrated model. (ten Tusscher et al., 2009) is a combined modeling and clinical study on the organization of ventricular fibrillation (VF) in the human heart. To support the credibility of the model used, epicardial excitation patterns are compared between model and clinical recordings obtained using a sock electrode, as is the time series of electrical activity at a fixed location. Dominant frequency of the time series is used for quantitative comparison between model and experiment. In addition, numbers of wavefronts and number of phase singularities over time are also compared. These quantities are convenient for condensing the complex spatio-temporal behavior of VF into simple time-series, useful for potential quantitative validation of very complex behavior.

Finally, we consider validation of organ-level CEP models with a specific clinical application; in particular, where a model is proposed to be used in clinical decision-making. One proposed application of CEP models is to use patient-specific simulations for risk stratification of patients with myocardial infarction, to determine which patients are at risk of SCD and therefore should receive prophylactic implantable cardioverter defibrillator (ICD) implant, as described in Arevalo et al. (2016). As illustrated in **Figure 5**, the software developed for this application uses patient-specific MR data to generate a biventricular mesh which includes regions of scar tissue and border zone. Electrical activity is simulated using the monodomain equations with the cell model of ten Tusscher et al. (2004). Various pacing protocols are virtually applied to determine if ventricular tachycardia (VT) is inducible, and if so the patient is classified as being at risk of SCD. We highlight two sets of validation results relevant to this model. The first, presented in Deng et al. (2015), is validation of epicardial excitation maps, for a swine version of the model, against swine data obtained using sock electrodes. The second, presented in Arevalo et al. (2016), describes a retrospective clinical study performed to test the risk index. In this study, the workflow described above was applied on a number of patients who had had ICD implant, and the risk classification as predicted by the model was compared to the clinical endpoint of ICD appropriately firing (or cardiac death). This is another form of (COU-driven) model validation, and of course it is a very strong form of validation because the QOI that is evaluated is the final QOI to be used in decision-making (i.e., risk index). Since it involves a clinical study, for this type of validation the appropriate quantitative analysis method is statistical; see (Arevalo et al., 2016) for details.

Other studies have proposed that related model-based tools could be used for prediction of ablation targets for patients affected by atrial fibrillation (AF) (McDowell et al., 2015), left atrial flutter (LAFL) (Zahid et al., 2016), or VT (Ashikaga et al., 2013). The proposed process for predicting ablation targets for all three arrhythmias is very similar to the process described above: an anatomically patient-specific model of the atria or ventricles is generated using MR data, and virtually interrogated to determine if AF/LAFL/VT is inducible. If so, ablation sites can be predicted that render AF/LAFL/VT non-inducible (a method for doing so is described in Zahid et al., 2016). (Zahid et al., 2016) presents a retrospective clinical study in which predicted LAFL ablation sites using the patient-specific atrial models are compared to the clinically ablated sites. (Chen et al., 2016) present a related workflow for developing patient-specific cardiac models, with personalisation of some functional parameters as well as anatomical personalisation. They perform validation by comparing model predictions of VT inducibility and re-entrant circuits to results from clinical VT stimulation studies.

Finally, similarly to cell models as discussed in section Cell Models, whole heart models have also been proposed to be used to predict drug-induced arrhythmogenic risk. (Okada et al., 2015) proposed that a heart and torso model which simulates the ECG can be used to integrate *in vitro* ion channel assays. The drug concentration at which Torsades de Pointes is induced in the simulated ECG is the proposed biomarker, and the predictive ability of this biomarker is tested using data for 12 drugs with well-characterized risk.

DISCUSSION

In this paper, we have categorized and discussed different types of evidence that could be used as a basis for the credibility of a CEP model. Our aim was to provide clarity on the body of evidence that typically is relevant (and often implied) in the evaluation of CEP models. As we transition into the era of Digital Health, there is a need for a systematic, rigorous and well-established methodology for justifying and assessing the credibility of computational models with biomedical applications. Current efforts toward these goals (ASME, 2017) are focused on "physics-based" models that have so far had most impact in medical devices applications (Pelton et al., 2008; Angelone et al., 2010; Stewart et al., 2012). However, these modeling fields are very different to physiological modeling in terms of model complexity, multi-scale nature, feasible validation, and inherent variability. In a previous publication we advocated that engineering model assessment methodologies of verification, validation and uncertainty quantification (VVUQ) could be used to improve credibility of models (Pathmanathan and Gray, 2013). However, while verification and uncertainty quantification methods are certainly relevant to CEP model assessment, best practices and quantitative methods in the engineering literature regarding validation appear less relevant to CEP models and other physiological models, due in part to the unique challenges in obtaining data for validation of physiological models. In general, the types of evidence supporting the credibility of physiological models will likely be very different to that for engineering models. Therefore, this paper is motivated by the need for a clear understanding of potential credibility evidence for CEP models, which can guide future efforts toward systematic
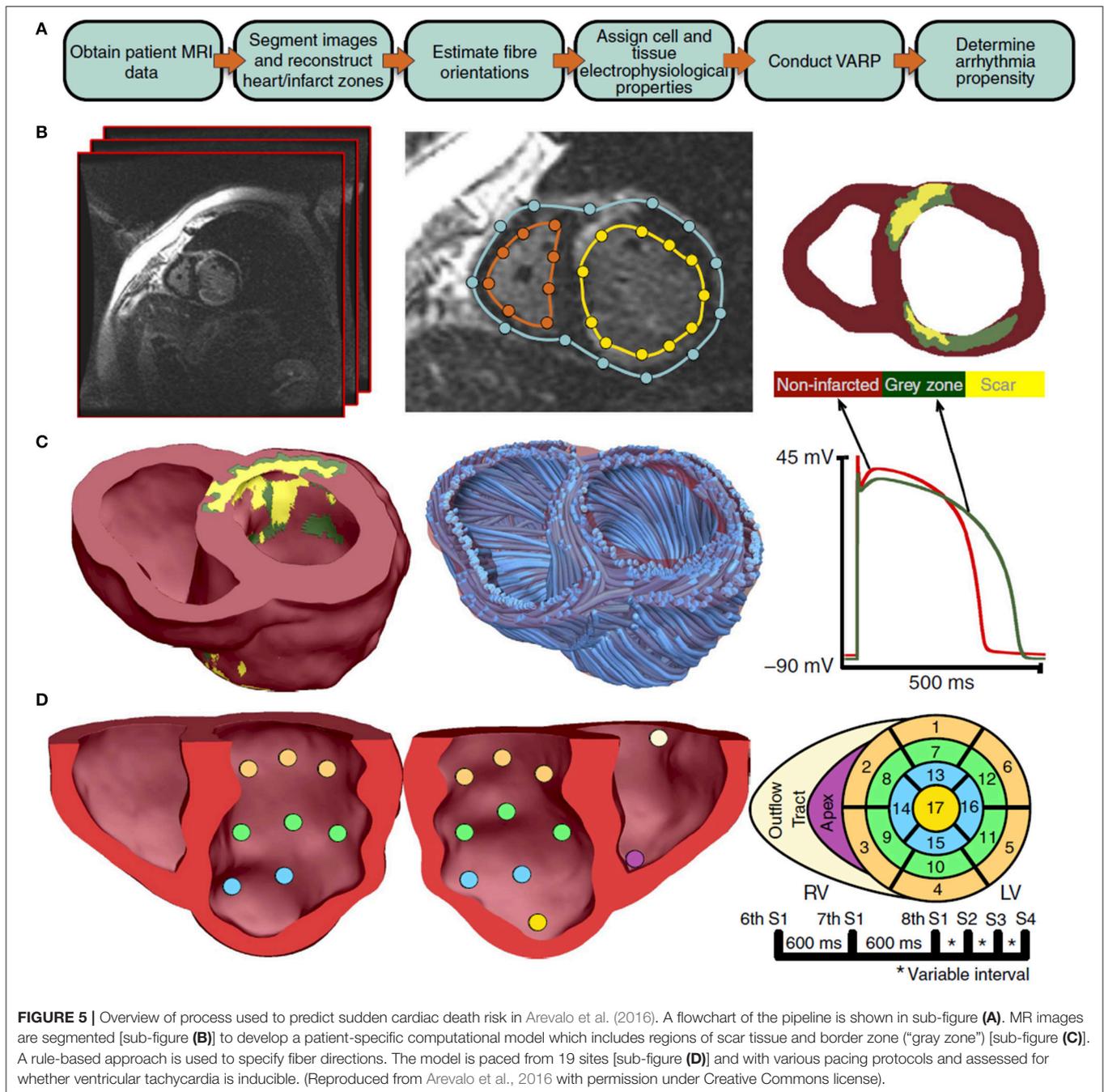
FIGURE 5 | Overview of process used to predict sudden cardiac death risk in Arevalo et al. (2016). A flowchart of the pipeline is shown in sub-figure **(A)**. MR images are segmented [sub-figure **(B)**] to develop a patient-specific computational model which includes regions of scar tissue and border zone ("gray zone") [sub-figure **(C)**]. A rule-based approach is used to specify fiber directions. The model is paced from 19 sites [sub-figure **(D)**] and with various pacing protocols and assessed for whether ventricular tachycardia is inducible. (Reproduced from Arevalo et al., 2016 with permission under Creative Commons license).

approaches for credibility assessment/justification which are relevant to physiological models.

We specifically highlighted validation of general-purpose CEP models not performed for any prescribed COU, and defined this as "general validation" evidence. As discussed in section Why Trust a Computational Model? the ever-increasing complexity of CEP cell models means that there are almost unlimited possibilities for such evaluation, and there is a large and ever-growing body of general validation results in the CEP modeling literature—in particular regarding cell models. Note that in this

review we described several examples of general validation but we did not discuss the "quality" of any general validation results. For example, we avoided subjective statements such as "validation results showed good agreement between model and experiment." This is because the level of agreement needed between model and experiment is determined by the COU, and when no COU is specified, a statement that a model shows "good agreement" without any context could potentially lead to inappropriate use of a model. In general, while general validation can provide important preliminary information about a computational

model, it may not be advisable to convert general validation results into binary "good"/"bad" or "acceptable"/"unacceptable" statements. However, when a COU of a model is chosen, previous general validation results can certainly be (re-)evaluated to determine how supportive they are of the model in the COU. This will likely require assessment of both the level of agreement between model and simulation, and also the relevance or "applicability" of the validation conditions to the COU; discussed in detail in Pathmanathan et al. (2017).

Currently, general validation results for cell models that are published in the literature are not collected, curated, or made available in one place. One resource that could potentially be useful for the cardiac modeling community is a resource on credibility evidence for cell models. The CEP modeling community already leads the way in model sharing and reproduction through the CellML repository and related software (Lloyd et al., 2008). The CellML language is a XML-based language for defining CEP cell (and other) models, allowing models to be defined unambiguously and easily shared, and the CellML repository serves as the starting point when using a cell model published in the literature. However, the repository does not include information regarding model validation results or other credibility evidence, and there is no way to easily look up such information. A sister repository containing model credibility evidence could therefore be useful to CEP model developers/users when deciding on which cell model to use for a particular COU. Examples of information that could be stored in such as repository include which emergent phenomena the model reproduces (and does not reproduce), and general validation results under a wide range of precisely prescribed protocols. One resource that provides a path toward such a repository is the Cardiac Electrophysiology Web Lab (Cooper et al., 2016). This is an online tool for easy comparison of multiple CellML-defined cell models under a wide range of protocols (which required the development of an XML-based language for specifying protocols Cooper et al., 2011). Being able to easily compare models is important because even models of the same species and heart region can behave quite differently; (see e.g., Cherry and Fenton, 2007). While the Web Lab does not currently provide explicit comparison to experimental data, it already serves as a potential tool for identifying which phenomena models can reproduce, and one can imagine an extension in which experimental data (from a wide range of sources and with full details on experimental conditions and protocols) are also included and comparison to model predictions are provided, both visually for qualitative comparison and perhaps quantitatively with appropriate validation metrics. In fact, inclusion of experimental data is one of the future plans of the Web Lab developers (personal communications). As stated above, we believe such results should not be converted into binary good/bad or acceptable/unacceptable judgements, or used to rank models. Instead, such a repository would serve as a rich resource by providing information needed for selecting between competing models for a particular COU, as well as providing validation results that could serve as a starting point for justification of model credibility for the COU. Moreover, if users were able to upload models and automatically run all protocols

(already possible in Web Lab) and then compare against the experimental data, this would be a powerful tool for validation of *modified* cell models (examples of which were provided in section Organ-Level Models), i.e., for comprehensive "re-validation." Note that we are *not* stating that an altered model should only be used if it "passes all validation tests." Indeed, for many COUs, a model not reproducing given phenomena could be argued to be acceptable given the COU. The point is that trust in cardiac models can be improved by collection of evidence, glass box evaluation, and explicit justification that the model is sufficiently credible for the COU despite its limitations.

It can be difficult to determine whether results presented in publications are calibration or validation results, as we mentioned in section Credibility of CEP Models at Different Spatial Scales. Specifically, while figures may be provided in which simulation and experimental results can be visually (and qualitatively) compared, it is often unclear whether any model parameters were tuned, optimized or tweaked to obtain the agreement with the experimental data. When that is the case, the results are calibration results, which is fundamentally weaker credibility evidence than validation of the *completed* model. Therefore, ideally calibration and validation results should be presented separately. While we believe that the examples of validation discussed in this paper are genuine validation results, it is certainly possible that some are actually calibration results. We also mentioned how simulation studies using CEP models are often performed in which no validation results or discussion of model credibility is presented. Such studies essentially implicitly rely—not unreasonably—on the maturity of the field and the various sources of historical evidence that we discussed in sections Ion Channel Models, Cell Models, and Organ-Level Models. The problem with this approach is it can contribute to a lack of clarity in the literature about the trustworthiness of CEP models, which can potentially lead to overconfidence in CEP models by non-experts who are unfamiliar with model subtleties (see initial discussion in Gong et al., 2017) as well as under-confidence in simulation-based conclusions by those who are skeptical of computational models in medicine. Such skepticism may be one of the biggest hurdles that needs to be overcome for computational models to achieve their potential in medical applications. These issues could be addressed by a clear and explicit presentation of the rationales for credibility of models used in simulation studies, referring as appropriate to the different sources of credibility evidence that support the use of the model for the COU as shown in **Table 1** (and/or appealing to the idea of models as representations, as discussed in section Why Trust a Computational Model?, when appropriate). As stated above, one aim is to argue that the model is sufficiently credible for the COU, despite model limitations. While there is no standardized method for determining what constitutes "sufficient credibility," the risk-informed strategy of (NASA, 2009; ASME, 2017) provides one method. The basic idea is that the credibility that needs to be demonstrated for a model should be related to the risk associated with incorrect predictions. Two factors are used to determine model risk. The first is model *influence*, which is the extent to which the model predictions will influence the decision to be made or conclusions of the study, compared

to other sources of information. The second is the *consequence* of incorrect decisions. For example, if a model is proposed to be used as the sole source of information in a safety-critical clinical decision, both influence and consequence are high, and the overall risk will be considered to be very high. Therefore, high credibility will be required of the model. In simulation studies, influence will often be high but consequence may be judged to be low, and overall risk may also be judged to be low, which means the credibility requirements are lower. Ultimately, we believe that routine and explicit justification of credibility will enable CEP models to have even greater impact in cardiac EP research, and facilitate their passage into clinical applications.

DISCLAIMER

The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services.

AUTHOR CONTRIBUTIONS

PP: Devised the paper, performed review of literature, and wrote the paper; RG: Provided feedback and edits on all aspects of paper.

REFERENCES

Aguado-Sierra, J., Krishnamurthy, A., Villongco, C., Chuang, J., Howard, E., Gonzales, M. J., et al. (2011). Patient-specific modeling of dyssynchronous heart failure: a case study. *Prog. Biophys. Mol. Biol.* 107, 147–155. doi: 10.1016/j.pbiomolbio.2011.06.014

Aguel, F., Eason, J., Trayanova, N., Iekas, G., and Fishler, M. (1999). Impact of transvenous lead position on active-Can ICD defibrillation: a computer simulation study. *Pacing Clin. Electrophysiol.* 22, 158–164. doi: 10.1111/j.1540-8159.1999.tb00324.x

Angelone, L. M., Ahveninen, J., Belliveau, J. W., and Bonmassar, G. (2010). Analysis of the role of lead resistivity in specific absorption rate for deep brain stimulator leads at 3T MRI. *IEEE Trans. Med. Imaging* 29, 1029–1038. doi: 10.1109/TMI.2010.2040624

Arevalo, H. J., Vadakkumpadan, F., Guallar, E., Jebb, A., Malamas, P., Wu, K. C., et al. (2016). Arrhythmia risk stratification of patients after myocardial infarction using personalized heart models. *Nat. Commun.* 7:11437. doi: 10.1038/ncomms11437

Armstrong, C. M. (2006). Na channel inactivation from open and closed states. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17991–17996. doi: 10.1073/pnas.0607603103

Ashikaga, H., Arevalo, H., Vadakkumpadan, F., Blake, R.C. III., Bayer, J.D., Nazarian, S., et al. (2013). Feasibility of image-based simulation to estimate ablation target in human ventricular arrhythmia. *Heart Rhythm* 10, 1109–1116. doi: 10.1016/j.hrthm.2013.04.015

ASME, S. (2017). *Draft V&V 40 - Standard for Verification and Validation in Computational Methods for Medical Devices*. (New York, NY: American Society of Mechanical Engineers).

Baillargeon, B., Rebelo, N., Fox, D.D., Taylor, R.L., and Kuhl, E. (2014). The living heart project: a robust and integrative simulator for human heart function. *Eur. J. Mech. A Solids* 48, 38–47. doi: 10.1016/j.euromechsol.2014.04.001

Bayer, J. D., Blake, R. C., Plank, G., and Trayanova, N. A. (2012). A novel rule-based algorithm for assigning myocardial fiber orientation to computational heart models. *Ann. Biomed. Eng.* 40, 2243–2254. doi: 10.1007/s10439-012-0593-5

Beattie, K.A., Hill, A.P., Bardenet, R., Cui, Y., Vandenberg, J.I., Gavaghan, D.J., et al. (2017). Sinusoidal voltage protocols for rapid characterization of ion channel kinetics. *bioRxiv*. doi: 10.1101/100677

Beaumont, J., Roberge, F., and Leon, L. (1993). On the interpretation of voltage-clamp data using the Hodgkin-Huxley model. *Math. Biosci.* 115, 65–101. doi: 10.1016/0025-5564(93)90047-E

Bellocchi, G., Rivington, M., Donatelli, M., and Matthews, K. (2010). Validation of biophysical models: issues and methodologies. A review. *Agron. Sustain. Dev.* 30, 109–130. doi: 10.1051/agro/2009001

Berecki, G., Wilders, R., De Jonge, B., Van Ginneken, A. C., and Verkerk, A. O. (2010). Re-evaluation of the action potential upstroke velocity as a measure of the Na+ current in cardiac myocytes at physiological conditions. *PLoS ONE* 5:e15772. doi: 10.1371/journal.pone.0015772

Bishop, M. J., and Plank, G. (2012). The role of fine-scale anatomical structure in the dynamics of reentry in computational models of the rabbit ventricles. *J. Physiol.* 590, 4515–4535. doi: 10.1113/jphysiol.2012.229062

Bishop, M. J., Hales, P., Plank, G., Gavaghan, D. J., Scheider, J., and Grau, V. (2009). "Comparison of Rule-Based and DTMRI-Derived Fibre Architecture in a whole rat ventricular computational model," in *Functional Imaging and Modeling of the Heart: 5th International Conference, FIMH 2009, Nice, France, Proceedings* eds. N. Ayache, H. Delingette, and M. Sermesant. (Berlin; Heidelberg: Springer Berlin Heidelberg), 87–96.

Bishop, M. J., Plank, G., Burton, R. A., Schneider, J. E., Gavaghan, D. J., Grau, V., et al. (2010). Development of an anatomically detailed MRI-derived rabbit ventricular model and assessment of its impact on simulations of electrophysiological function. *Am. J. Physiol. Heart Circ. Physiol.* 298, H699–H718. doi: 10.1152/ajpheart.00606.2009

Bishop, M. J., Rodriguez, B., Qu, F., Efimov, I. R., Gavaghan, D. J., and Trayanova, N. A. (2007). The role of photon scattering in optical signal distortion during arrhythmia and defibrillation. *Biophys. J.* 93, 3714–3726. doi: 10.1529/biophysj.107.110981

Bordas, R., Gillow, K., Lou, Q., Efimov, I., Gavaghan, D., Kohl, P., et al. (2011). Rabbit-specific ventricular model of cardiac electrophysiological function including specialized conduction system. *Prog. Biophys. Mol. Biol.* 107, 90–100. doi: 10.1016/j.pbiomolbio.2011.05.002

Bourgault, Y., Coudiere, Y., and Pierre, C. (2009). Existence and uniqueness of the solution for the bidomain model used in cardiac electrophysiology. *Nonlinear Anal. Real World Appl.* 10, 458–482. doi: 10.1016/j.nonrwa.2007.10.007

Brodland, G. W. (2015). How computational models can help unlock biological systems. *Semin. Cell Dev. Biol.* 47, 62–73. doi: 10.1016/j.semcdb.2015.07.001

Bueno-Orovio, A., Kay, D., Grau, V., Rodriguez, B., and Burrage, K. (2014). Fractional diffusion models of cardiac electrical propagation: role of structural heterogeneity in dispersion of repolarization. *J. R. Soc. Interface* 11:20140352. doi: 10.1098/rsif.2014.0352

Carro, J., Rodríguez-Matas, J. F., Monasterio, V., and Pueyo, E. (2017). Limitations in electrophysiological model development and validation caused by differences between simulations and experimental protocols. *Prog. Biophys. Mol. Biol.* 129, 53–64. doi: 10.1016/j.pbiomolbio.2016.11.006

Carusi, A., Burrage, K., and Rodríguez, B. (2012). Bridging experiments, models and simulations: an integrative approach to validation in computational cardiac electrophysiology. *Am. J. Physiol. Heart Circ. Physiol.* 303, H144–H155. doi: 10.1152/ajpheart.01151.2011

Cavero, I., and Holzgrefe, H. (2014). Comprehensive *in vitro* proarrhythmia assay, a novel *in vitro/in silico* paradigm to detect ventricular proarrhythmic liability: a visionary 21st century initiative. *Expert Opin. Drug Saf.* 13, 745–758. doi: 10.1517/14740338.2014.915311

Chabiniok, R., Wang, V. Y., Hadjicharalambous, M., Asner, L., Lee, J., Sermesant, M., et al. (2016). Multiphysics and multiscale modelling, data–model fusion and integration of organ physiology in the clinic: ventricular cardiac mechanics. *Interface Focus* 6:20150083. doi: 10.1098/rsfs.2015.0083

Chang, E. T., Strong, M., and Clayton, R. H. (2015). Bayesian sensitivity analysis of a cardiac cell model using a Gaussian process emulator. *PLoS ONE* 10:e0130252. doi: 10.1371/journal.pone.0130252

Chang, K. C., Dutta, S., Mirams, G. R., Beattie, K. A., Sheng, J., Tran, P. N., et al. (2017). Uncertainty quantification reveals the importance of data

variability and experimental design considerations for *in silico* proarrhythmia risk assessment. *Front. Physiol.* 8:917. doi: 10.3389/fphys.2017.00917

Chen, Z., Cabrera-Lozoya, R., Relan, J., Sohal, M., Shetty, A., Karim, R., et al. (2016). Biophysical modeling predicts ventricular tachycardia inducibility and circuit morphology: a combined clinical validation and computer modeling approach. *J. Cardiovasc. Electrophysiol.* 27, 851–860. doi: 10.1111/jce.12991

Cherry, E. M., and Fenton, F. H. (2007). A tale of two dogs: analyzing two models of canine ventricular electrophysiology. *Am. J. Physiol. Heart Circ. Physiol.* 292, H43–H55. doi: 10.1152/ajpheart.00955.2006

Clancy, C. E., and Rudy, Y. (1999). Linking a genetic defect to its cellular phenotype in a cardiac arrhythmia. *Nature* 400, 566. doi: 10.1038/23034

Clayton, R., Bernus, O., Cherry, E., Dierckx, H., Fenton, F., Mirabella, L., et al. (2011). Models of cardiac tissue electrophysiology: progress, challenges and open questions. *Prog. Biophys. Mol. Biol.* 104, 22–48. doi: 10.1016/j.pbiomolbio.2010.05.008

Colatsky, T., Fermini, B., Gintant, G., Pierson, J. B., Sager, P., Sekino, Y., et al. (2016). The comprehensive *in vitro* proarrhythmia assay (CiPA) initiative—Update on progress. *J. Pharmacol. Toxicol. Methods* 81, 15–20. doi: 10.1016/j.vascn.2016.06.002

Cole, K. S., and Moore, J. W. (1960). Ionic current measurements in the squid giant axon membrane. *J. Gen. Physiol.* 44, 123–167. doi: 10.1085/jgp.44.1.123

Connolly, A., Vigmond, E., and Bishop, M. (2017). Virtual electrodes around anatomical structures and their roles in defibrillation. *PLoS ONE* 12:e0173324. doi: 10.1371/journal.pone.0173324

Cooper, J., Mirams, G. R., and Niederer, S. A. (2011). High-throughput functional curation of cellular electrophysiology models. *Prog. Biophys. Mol. Biol.* 107, 11–20. doi: 10.1016/j.pbiomolbio.2011.06.003

Cooper, J., Scharm, M., and Mirams, G. R. (2016). The cardiac electrophysiology web lab. *Biophys. J.* 110, 292–300. doi: 10.1016/j.bpj.2015.12.012

Csercsik, D., Hangos, K. M., and Szederkényi, G. (2012). Identifiability analysis and parameter estimation of a single Hodgkin–Huxley type voltage dependent ion channel under voltage step measurement conditions. *Neurocomputing* 77, 178–188. doi: 10.1016/j.neucom.2011.09.006

Davies, M. R., Mistry, H. B., Hussein, L., Pollard, C. E., Valentin, J.-P., Swinton, J., et al. (2012). An *in silico* canine cardiac midmyocardial action potential duration model as a tool for early drug safety assessment. *Am. J. Physiol. Heart Circ. Physiol.* 303, H497–H498. doi: 10.1152/ajpheart.00808.2011

Davies, M. R., Wang, K., Mirams, G. R., Caruso, A., Noble, D., Walz, A., et al. (2016). Recent developments in using mechanistic cardiac modelling for drug safety evaluation. *Drug Discov. Today* 21, 924–938. doi: 10.1016/j.drudis.2016.02.003

Deng, D., Arevalo, H., Pashakhanloo, F., Prakosa, A., Ashikaga, H., McVeigh, E., et al. (2015). Accuracy of prediction of infarct-related arrhythmic circuits from image-based models reconstructed from low and high resolution MRI. *Front. Physiol.* 6:282. doi: 10.3389/fphys.2015.00282

Deng, D., Jiao, P., Ye, X., and Xia, L. (2012). An image-based model of the whole human heart with detailed anatomical structure and fiber orientation. *Comput. Math. Methods Med.* 2012:891070. doi: 10.1155/2012/891070

Dierckx, H., Benson, A. P., Gilbert, S. H., Ries, M. E., Holden, A. V., Verschelde, H., et al. (2009). "Intravoxel fibre structure of the left ventricular free wall and posterior left-right ventricular insertion site in canine myocardium using Q-Ball imaging," in *International Conference on Functional Imaging and Modeling of the Heart* (Springer), 495–504. doi: 10.1007/978-3-642-01932-6_53

Dillon, S. M., Kerner, T. E., Hoffman, J., Menz, V., Li, K., and Michele, J. (1998). A system for *in-vivo* cardiac optical mapping. *IEEE Eng. Med. Biol. Mag.* 17, 95–108. doi: 10.1109/51.646226

Dutta, S., Chang, K. C., Beattie, K. A., Sheng, J., Tran, P. N., Wu, W. W., et al. (2017). Optimization of an *in silico* cardiac cell model for proarrhythmia risk assessment. *Front. Physiol.* 8:616. doi: 10.3389/fphys.2017.00616

Eddy, D. M., Hollingworth, W., Caro, J. J., Tsevat, J., McDonald, K. M., and Wong, J. B. (2012). Model transparency and validation a report of the ISPOR-SMDM modeling good research practices task force—7. *Med. Decis. Making* 32, 733–743. doi: 10.1177/0272989X12454579

Elkins, R.C., Davies, M.R., Brough, S.J., Gavaghan, D.J., Cui, Y., Abi-Gerges, N., et al. (2013). Variability in high-throughput ion-channel screening data and consequences for cardiac safety assessment. *J. Pharmacol. Toxicol. Methods* 68, 112–122. doi: 10.1016/j.vascn.2013.04.007

Fink, M., Niederer, S. A., Cherry, E. M., Fenton, F. H., Koivumäki, J. T., Seemann, G., et al. (2011). Cardiac cell modelling: observations from the heart of the cardiac physiome project. *Prog. Biophys. Mol. Biol.* 104, 2–21. doi: 10.1016/j.pbiomolbio.2010.03.002

Food and Drug Administration, (2016). *Reporting of Computational Modeling Studies in Medical Device Submissions—Guidance for Industry and Food and Drug Administration Staff.* (Washington, DC: Food and Drug Administration).

Franzone, P. C., Pavarino, L. F., and Scacchi, S. (2014). *Mathematical Cardiac Electrophysiology.* New York, NY: Springer.

Gong, J. Q., Shim, J. V., Nú-ez-Acosta, E., and Sobie, E. A. (2017). I love it when a plan comes together: insight gained through convergence of competing mathematical models. *J. Mol. Cell. Cardiol.* 102, 31–33. doi: 10.1016/j.yjmcc.2016.10.015

Grandi, E., Morotti, S., Ginsburg, K. S., Severi, S., and Bers, D. M. (2010). Interplay of voltage and Ca-dependent inactivation of L-type Ca current. *Prog. Biophys. Mol. Biol.* 103, 44–50. doi: 10.1016/j.pbiomolbio.2010.02.001

Gray, R. A. (1999). What exactly are optically recorded "action potentials"? *J. Cardiovasc. Electrophysiol.* 10, 1463–1466. doi: 10.1111/j.1540-8167.1999.tb00205.x

Gray, R. A., Mashburn, D. N., Sidorov, V. Y., and Wikswo, J. P. (2013). Quantification of transmembrane currents during action potential propagation in the heart. *Biophys. J.* 104, 268–278. doi: 10.1016/j.bpj.2012.11.007

Häusser, M. (2000). The Hodgkin-Huxley theory of the action potential. *Nat. Neurosci.* 3, 1165–1165. doi: 10.1038/81426

Henriquez, C. S. (1993). Simulating the electrical behavior of cardiac tissue using the bidomain model. *Crit. Rev. Biomed. Eng.* 21, 1–77.

Hills, R. G., Pilch, M., Dowding, K. J., Red-Horse, J., Paez, T. L., Babuška, I., et al. (2008). Validation challenge workshop. *Comput. Methods Appl. Mech. Eng.* 197, 2375–2380. doi: 10.1016/j.cma.2007.10.016

Hodgkin, A. L., and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* 117, 500–544. doi: 10.1113/jphysiol.1952.sp004764

Hyatt, C. J., Mironov, S. F., Wellner, M., Berenfeld, O., Popp, A. K., Weitz, D. A., et al. (2003). Synthesis of voltage-sensitive fluorescence signals from three-dimensional myocardial activation patterns. *Biophys. J.* 85, 2673–2683. doi: 10.1016/S0006-3495(03)74690-6

Iyer, V., Mazhari, R., and Winslow, R.L. (2004). A computational model of the human left-ventricular epicardial myocyte. *Biophys. J.* 87, 1507–1525. doi: 10.1529/biophysj.104.043299

Jacquez, J. A., and Greif, P. (1985). Numerical parameter identifiability and estimability: integrating identifiability, estimability, and optimal sampling design. *Math. Biosci.* 77, 201–227. doi: 10.1016/0025-5564(85)90098-7

Johnstone, R. H., Chang, E. T., Bardenet, R., De Boer, T. P., Gavaghan, D. J., Pathmanathan, P., et al. (2016). Uncertainty and variability in models of the cardiac action potential: can we build trustworthy models? *J. Mol. Cell. Cardiol.* 96, 49–62. doi: 10.1016/j.yjmcc.2015.11.018

Keldermann, R. H., Ten Tusscher, K. H. W. J., Nash, M. P., Hren, R., Taggart, P., and Panfilov, A. V. (2008). Effect of heterogeneous APD restitution on VF organization in a model of the human ventricles. *Am. J. Physiol. Heart Circ. Physiol.* 294, H764–H774. doi: 10.1152/ajpheart.00906.2007

Konukoglu, E., Relan, J., Cilingir, U., Menze, B. H., Chinchapatnam, P., Jadidi, A., et al. (2011). Efficient probabilistic model personalization integrating uncertainty on data and parameters: application to eikonal-diffusion models in cardiac electrophysiology. *Prog. Biophys. Mol. Biol.* 107, 134–146. doi: 10.1016/j.pbiomolbio.2011.07.002

Krishnamoorthi, S., Perotti, L. E., Borgstrom, N. P., Ajijola, O. A., Frid, A., Ponnaluri, A. V., et al. (2014). Simulation methods and validation criteria for modeling cardiac ventricular electrophysiology. *PLoS ONE* 9:e114494. doi: 10.1371/journal.pone.0114494

Krueger, M. W., Schmidt, V., Tobón, C., Weber, F. M., Lorenz, C., Keller, D. U. J., et al. (2011). "Modeling atrial fiber orientation in patient-specific geometries: a semi-automatic rule-based approach," in *Functional Imaging and Modeling of the Heart: 6th International Conference, FIMH 2011, New York City, NY, USA Proceedings,* eds D. N. Metaxas and L. Axel (Berlin; Heidelberg: Springer Berlin Heidelberg), 223–232. doi: 10.1007/978-3-642-21028-0_28

Lee, J., Smaill, B., and Smith, N. (2006). Hodgkin-Huxley type ion channel characterization: an improved method of voltage clamp experiment parameter estimation. *J. Theor. Biol.* 242, 123–134. doi: 10.1016/j.jtbi.2006.02.006

Legrice, I. J., Smaill, B., Chai, L., Edgar, S., Gavin, J., and Hunter, P. J. (1995). Laminar structure of the heart: ventricular myocyte arrangement and connective tissue architecture in the dog. *Am. J. Physiol. Heart Circ. Physiol.* 269, H571–H582. doi: 10.1152/ajpheart.1995.269.2.H571

Lloyd, C. M., Lawson, J. R., Hunter, P. J., and Nielsen, P. F. (2008). The CellML model repository. *Bioinformatics* 24, 2122–2123. doi: 10.1093/bioinformatics/btn390

Lopez-Perez, A., Sebastian, R., and Ferrero, J. M. (2015). Three-dimensional cardiac computational modelling: methods, features and applications. *Biomed. Eng. Online* 14:35. doi: 10.1186/s12938-015-0033-5

Mann, S. A., Imtiaz, M., Winbo, A., Rydberg, A., Perry, M. D., Couderc, J.-P., et al. (2016). Convergence of models of human ventricular myocyte electrophysiology after global optimization to recapitulate clinical long QT phenotypes. *J. Mol. Cell. Cardiol.* 100, 25–34. doi: 10.1016/j.yjmcc.2016.09.011

Mauro, A., Conti, F., Dodge, F., and Schor, R. (1970). Subthreshold behavior and phenomenological impedance of the squid giant axon. *J. Gen. Physiol.* 55, 497–523. doi: 10.1085/jgp.55.4.497

McDowell, K. S., Vadakkumpadan, F., Blake, R., Blauer, J., Plank, G., Macleod, R. S., et al. (2012). Methodology for patient-specific modeling of atrial fibrosis as a substrate for atrial fibrillation. *J. Electrocardiol.* 45, 640–645. doi: 10.1016/j.jelectrocard.2012.08.005

McDowell, K. S., Zahid, S., Vadakkumpadan, F., Blauer, J., Macleod, R. S., and Trayanova, N. A. (2015). Virtual electrophysiological study of atrial fibrillation in fibrotic remodeling. *PLoS ONE* 10:e0117110. doi: 10.1371/journal.pone.0117110

Mekkaoui, C., Huang, S., Chen, H. H., Dai, G., Reese, T. G., Kostis, W. J., et al. (2012). Fiber architecture in remodeled myocardium revealed with a quantitative diffusion CMR tractography framework and histological validation. *J. Cardiovasc. Magn. Reson.* 14:70. doi: 10.1186/1532-429X-14-70

Mirams, G. R., Cui, Y., Sher, A., Fink, M., Cooper, J., Heath, B. M., et al. (2011). Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovasc. Res.* 91, 53–61. doi: 10.1093/cvr/cvr044

Mirams, G. R., Pathmanathan, P., Gray, R. A., Challenor, P., and Clayton, R. H. (2016). Uncertainty and variability in computational and mathematical models of cardiac physiology. *J. Physiol.* 594, 6833–6847. doi: 10.1113/JP271671

Muzikant, A., and Henriquez, C. (1998). Validation of three-dimensional conduction models using experimental mapping: are we getting closer? *Prog. Biophys. Mol. Biol.* 69, 205–223. doi: 10.1016/S0079-6107(98)00008-X

Muzikant, A. L., Hsu, E. W., Wolf, P. D., and Henriquez, C. S. (2002). Region specific modeling of cardiac muscle: comparison of simulated and experimental potentials. *Ann. Biomed. Eng.* 30, 867–883. doi: 10.1114/1.1509453

NASA (2009). *7009: Standard for Models and Simulation*. Washington, DC: NASA.

National Research Council (2012). *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*. Washington, DC: National Academies Press.

Nesterenko, V. V., Zygmunt, A. C., Rajamani, S., Belardinelli, L., and Antzelevitch, C. (2011). Mechanisms of atrial-selective block of Na+ channels by ranolazine: II. Insights from a mathematical model. *Am. J. Physiol. Heart Circ. Physiol.* 301, H1615–H1624. doi: 10.1152/ajpheart.00243.2011

Neu, J., and Krassowska, W. (1993). Homogenization of syncytial tissues. *Crit. Rev. Biomed. Eng.* 21, 137–199.

Niederer, S. A., and Smith, N. P. (2016). Using physiologically based models for clinical translation: predictive modelling, data interpretation or something in-between? *J. Physiol.* 594, 6849–6863. doi: 10.1113/JP272003

Niederer, S. A., Kerfoot, E., Benson, A. P., Bernabeu, M. O., Bernus, O., Bradley, C., et al. (2011). Verification of cardiac tissue electrophysiology simulators using an N-version benchmark. *Phil. Trans. R. Soc. A* 369, 4331–4351. doi: 10.1098/rsta.2011.0139

Niederer, S. A., Plank, G., Chinchapatnam, P., Ginks, M., Lamata, P., Rhode, K. S., et al. (2010). Length-dependent tension in the failing heart and the efficacy of cardiac resynchronization therapy. *Cardiovasc. Res.* 89, 336–343. doi: 10.1093/cvr/cvq318

Noble, D. (1962). A modification of the Hodgkin—Huxley equations applicable to Purkinje fibre action and pacemaker potentials. *J. Physiol.* 160, 317–352. doi: 10.1113/jphysiol.1962.sp006849

Noble, D. (2011). Successes and failures in modeling heart cell electrophysiology. *Heart Rhythm* 8, 1798–1803. doi: 10.1016/j.hrthm.2011.06.014

Noble, D., Garny, A., and Noble, P. J. (2012). How the Hodgkin–Huxley equations inspired the cardiac physiome project. *J. Physiol.* 590, 2613–2628. doi: 10.1113/jphysiol.2011.224238

Nordin, C., and Ming, Z. (1995). Computer model of current-induced early afterdepolarizations in guinea pig ventricular myocytes. *Am. J. Physiol. Heart Circ. Physiol.* 268, H2440–H2459. doi: 10.1152/ajpheart.1995.268.6.H2440

Oberkampf, W. L., and Barone, M. F. (2006). Measures of agreement between computation and experiment: validation metrics. *J. Comput. Phys.* 217, 5–36. doi: 10.1016/j.jcp.2006.03.037

Oberkampf, W. L., Trucano, T. G., and Hirsch, C. (2004). Verification, validation, and predictive capability in computational engineering and physics. *Appl. Mech. Rev.* 57:345. doi: 10.1115/1.1767847

O'Hara, T., Virág, L., Varró, A., and Rudy, Y. (2011). Simulation of the undiseased human cardiac ventricular action potential: model formulation and experimental validation. *PLoS Comput. Biol.* 7:e1002061. doi: 10.1371/journal.pcbi.1002061

Okada, J.-I., Yoshinaga, T., Kurokawa, J., Washio, T., Furukawa, T., Sawada, K., et al. (2015). Screening system for drug-induced arrhythmogenic risk combining a patch clamp and heart simulator. *Sci. Adv.* 1:e1400142. doi: 10.1126/sciadv.1400142

Passini, E., Britton, O. J., Lu, H. R., Rohrbacher, J., Hermans, A. N., Gallacher, D. J., et al. (2017). Human *in silico* drug trials demonstrate higher accuracy than animal models in predicting clinical pro-arrhythmic cardiotoxicity. *Front. Physiol.* 8:668. doi: 10.3389/fphys.2017.00668

Pathmanathan, P., and Gray, R. A. (2013). Ensuring reliability of safety-critical clinical applications of computational cardiac models. *Front. Physiol.* 4:358. doi: 10.3389/fphys.2013.00358

Pathmanathan, P., and Gray, R. A. (2014). Verification of computational models of cardiac electro-physiology. *Int. J. Numer. Methods Biomed. Eng.* 30, 525–544. doi: 10.1002/cnm.2615

Pathmanathan, P., Gray, R. A., Romero, V. J., and Morrison, T. M. (2017). Applicability analysis of validation evidence for biomedical computational models. *J. Verif. Valid. Uncert. Quantification* 2, 021005–021005-11. doi: 10.1115/1.4037671

Pathmanathan, P., Shotwell, M. S., Gavaghan, D. J., Cordeiro, J. M., and Gray, R. A. (2015). Uncertainty quantification of fast sodium current steady-state inactivation for multi-scale models of cardiac electrophysiology. *Prog. Biophys. Mol. Biol.* 117, 4–18. doi: 10.1016/j.pbiomolbio.2015.01.008

Patterson, E. A., and Whelan, M. P. (2017). A framework to establish credibility of computational models in biology. *Prog. Biophys. Mol. Biol.* 129, 13–19. doi: 10.1016/j.pbiomolbio.2016.08.007

Pelton, A., Schroeder, V., Mitchell, M., Gong, X.-Y., Barney, M., and Robertson, S. (2008). Fatigue and durability of Nitinol stents. *J. Mech. Behav. Biomed. Mater.* 1, 153–164. doi: 10.1016/j.jmbbm.2007.08.001

Plank, G., Burton, R. A., Hales, P., Bishop, M., Mansoori, T., Bernabeu, M. O., et al. (2009). Generation of histo-anatomically representative models of the individual heart: tools and application. *Philos. Trans. R. Soc. Lond. A Math. Phys. Eng. Sci.* 367, 2257–2292. doi: 10.1098/rsta.2009.0056

Potse, M., Dubé, B., Richer, J., Vinet, A., and Gulrajani, R. M. (2006). A comparison of monodomain and bidomain reaction-diffusion models for action potential propagation in the human heart. *IEEE Trans. Biomed. Eng.* 53, 2425–2435. doi: 10.1109/TBME.2006.880875

Raba, A. E., Cordeiro, J. M., Antzelevitch, C., and Beaumont, J. (2013). Extending the conditions of application of an inversion of the Hodgkin–Huxley Gating model. *Bull. Math. Biol.* 75, 752–773. doi: 10.1007/s11538-013-9832-7

Relan, J., Pop, M., Delingette, H., Wright, G. A., Ayache, N., and Sermesant, M. (2011). Personalization of a cardiac electrophysiology model using optical mapping and MRI for prediction of changes with pacing. *IEEE Trans. Biomed. Eng.* 58, 3339–3349. doi: 10.1109/TBME.2011.2107513

Richards, D. F., Glosli, J. N., Draeger, E. W., Mirin, A. A., Chan, B., Fattebert, J.-L., et al. (2013). Towards real-time simulation of cardiac electrophysiology in a human heart at high resolution. *Comput. Methods Biomech. Biomed. Eng.* 16, 802–805. doi: 10.1080/10255842.2013.795556

Richardson, G., and Chapman, S. J. (2011). Derivation of the bidomain equations for a beating heart with a general microstructure. *SIAM J. Appl. Math.* 71, 657–675. doi: 10.1137/090777165

Roache, P. J. (2009). Perspective: validation—What does it mean? *J. Fluids Eng.* 131:034503. doi: 10.1115/1.3077134

Rodriguez, B., Li, L., Eason, J. C., Efimov, I. R., and Trayanova, N. A. (2005). Differences between left and right ventricular chamber geometry affect cardiac vulnerability to electric shocks. *Circ. Res.* 97, 168–175. doi: 10.1161/01.RES.0000174429.00987.17

Romero, D., Sebastian, R., Bijnens, B. H., Zimmerman, V., Boyle, P. M., Vigmond, E. J., et al. (2010). Effects of the purkinje system and cardiac geometry on biventricular pacing: a model study. *Ann. Biomed. Eng.* 38, 1388–1398. doi: 10.1007/s10439-010-9926-4

Rossi, S., and Griffith, B. E. (2017). Incorporating inductances in tissue-scale models of cardiac electrophysiology. *Chaos* 27, 093926. doi: 10.1063/1.5000706

Roth, B. J., and Pertsov, A. M. (2009). Hybrid modeling of electrical and optical behavior in the heart. *Physica D Nonlinear Phenomena* 238, 1019–1027. doi: 10.1016/j.physd.2009.01.014

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica* 577–591. doi: 10.2307/1913267

Schwiening, C. J. (2012). A brief historical perspective: Hodgkin and Huxley. *J. Physiol.* 590, 2571–2575. doi: 10.1113/jphysiol.2012.230458

Seemann, G., Höper, C., Sachse, F. B., Dössel, O., Holden, A. V., and Zhang, H. (2006). Heterogeneous three-dimensional anatomical and electrophysiological model of human atria. *Philos. Trans. R. Soc. Lond. A Math. Phys. Eng. Sci.* 364, 1465–1481. doi: 10.1098/rsta.2006.1781

Sepulveda, N. G., Roth, B. J., and Wikswo, J. (1989). Current injection into a two-dimensional anisotropic bidomain. *Biophys. J.* 55, 987–999. doi: 10.1016/S0006-3495(89)82897-8

Shotwell, M. S., and Gray, R. A. (2016). Estimability analysis and optimal design in dynamic multi-scale models of cardiac electrophysiology. *J. Agric. Biol. Environ. Stat.* 21, 261–276. doi: 10.1007/s13253-016-0244-7

Smith, N., De Vecchi, A., McCormick, M., Nordsletten, D., Camara, O., Frangi, A. F., et al. (2011). euHeart: personalized and integrated cardiac care using patient-specific cardiovascular modelling. *Interface Focus* 1, 349–364. doi: 10.1098/rsfs.2010.0048

Smith, R. C. (2013). *Uncertainty Quantification: Theory, Implementation, and Applications.* Philadelphia, PA: SIAM.

Sobie, E. A. (2009). Parameter sensitivity analysis in electrophysiological models using multivariable regression. *Biophys. J.* 96, 1264–1274. doi: 10.1016/j.bpj.2008.10.056

Stewart, S. F., Paterson, E. G., Burgreen, G. W., Hariharan, P., Giarra, M., Reddy, V., et al. (2012). Assessment of CFD performance in simulations of an idealized medical device: results of FDA's first computational interlaboratory study. *Cardiovasc. Eng. Technol.* 3, 139–160. doi: 10.1007/s13239-012-0087-5

ten Tusscher, K. H., and Panfilov, A. V. (2006). Cell model for efficient simulation of wave propagation in human ventricular tissue under normal and pathological conditions. *Phys. Med. Biol.* 51, 6141–6156. doi: 10.1088/0031-9155/51/23/014

ten Tusscher, K. H., Mourad, A., Nash, M., Clayton, R. H., Bradley, C. P., Paterson, D. J., et al. (2009). Organization of ventricular fibrillation in the human heart: experiments and models. *Exp. Physiol.* 94, 553–562. doi: 10.1113/expphysiol.2008.044065

ten Tusscher, K. H., Noble, D., Noble, P. J., and Panfilov, A. V. (2004). A model for human ventricular tissue. *Am. J. Physiol. Heart Circ. Physiol.* 286, H1573–H1589. doi: 10.1152/ajpheart.00794.2003

Trayanova, N., Plank, G., and Rodríguez, B. (2006). What have we learned from mathematical models of defibrillation and postshock arrhythmogenesis? Application of bidomain simulations. *Heart Rhythm* 3, 1232. doi: 10.1016/j.hrthm.2006.04.015

Trayanova, N. A. (2011). Whole-heart modeling. *Circ. Res.* 108, 113–128. doi: 10.1161/CIRCRESAHA.110.223610

Trayanova, N. A., and Boyle, P. M. (2014). Advances in modeling ventricular arrhythmias: from mechanisms to the clinic. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 6, 209–224. doi: 10.1002/wsbm.1256

Tung, L. (1978). *A bi-Domain Model for Describing Ischemic Myocardial dc Potentials.* Massachusetts Institute of Technology.

Uzelac, I., Ji, Y.C., Hornung, D., Schröder-Scheteling, J., Luther, S., Gray, R. A., et al. (2017). Simultaneous quantification of spatially discordant alternans in voltage and intracellular calcium in langendorff-perfused rabbit hearts and inconsistencies with models of cardiac action potentials and Ca transients. *Front. Physiol.* 8:819. doi: 10.3389/fphys.2017.00819

Vetter, F. J., and McCulloch, A. D. (1998). Three-dimensional analysis of regional cardiac function: a model of rabbit ventricular anatomy. *Prog. Biophys. Mol. Biol.* 69, 157–183. doi: 10.1016/S0079-6107(98)00006-6

Viceconti, M. (2011). A tentative taxonomy for predictive models in relation to their falsifiability. *Phil. Trans. R. Soc. A* 369, 4149–4161. doi: 10.1098/rsta.2011.0227

Walter, E., and Pronzato, L. (1996). On the identifiability and distinguishability of nonlinear parametric models. *Math. Comput. Simul.* 42, 125–134. doi: 10.1016/0378-4754(95)00123-9

Wang, G. J., and Beaumont, J. (2004). Parameter Estimation of the Hodgkin–Huxley gating model: an inversion procedure. *SIAM J. Appl. Math.* 64, 1249–1267. doi: 10.1137/S0036139902419826

Wikswo, J. P., Lin, S.-F., and Abbas, R. A. (1995). Virtual electrodes in cardiac tissue: a common mechanism for anodal and cathodal stimulation. *Biophys. J.* 69, 2195–2210. doi: 10.1016/S0006-3495(95)80115-3

Wikswo, J. P., and Roth, B. J. (2009). "Virtual electrode theory of pacing," in *Cardiac Bioelectric Therapy,* eds I. R. Efimov, M. W. Kroll, and P. J. Tchou (Boston, MA: Springer), 283–330.

Wikswo, J. P., Wisialowski, T. A., Altemeier, W. A., Balser, J. R., Kopelman, H. A., and Roden, D. M. (1991). Virtual cathode effects during stimulation of cardiac muscle. Two-dimensional *in vivo* experiments. *Circulation Res.* 68, 513–530. doi: 10.1161/01.RES.68.2.513

Winslow, R. L., Trayanova, N., Geman, D., and Miller, M. I. (2012). Computational medicine: translating models to clinical care. *Sci. Transl. Med.* 4, 158rv111. doi: 10.1126/scitranslmed.3003528

Yang, P. C., Song, Y., Giles, W. R., Horvath, B., Chen-Izu, Y., Belardinelli, L., et al. (2015). A computational modelling approach combined with cellular electrophysiology data provides insights into the therapeutic benefit of targeting the late Na+ current. *J. Physiol.* 593, 1429–1442. doi: 10.1113/jphysiol.2014.279554

Zahid, S., Whyte, K. N., Schwarz, E. L., Blake, R. C., Boyle, P. M., Chrispin, J., et al. (2016). Feasibility of using patient-specific models and the "minimum cut" algorithm to predict optimal ablation targets for left atrial flutter. *Heart Rhythm* 13, 1687–1698. doi: 10.1016/j.hrthm.2016.04.009

Zemzemi, N., Bernabeu, M. O., Saiz, J., Cooper, J., Pathmanathan, P., Mirams, G. R., et al. (2013). Computational assessment of drug-induced effects on the electrocardiogram: from ion channel to body surface potentials. *Br. J. Pharmacol.* 168, 718–733. doi: 10.1111/j.1476-5381.2012.02200.x

Zhao, J., Butters, T. D., Zhang, H., Legrice, I. J., Sands, G. B., and Smaill, B. H. (2013). Image-based model of atrial anatomy and electrical activation: a computational platform for investigating atrial arrhythmia. *IEEE Trans. Med. Imaging* 32, 18–27. doi: 10.1109/TMI.2012.2227776