Virtual Clinical Trials: Why and What (Special Section Guest Editorial)

Ehsan Samei,[a] **Paul Kinahan,**[b] **Robert M. Nishikawa,**[c] **and Andrew Maidment**[d]

[a]Duke University, Departments of Radiology, Physics, BME, and ECE, Durham, North Carolina, United States

[b]University of Washington, Departments of Radiology and Bioengineering, Seattle, Washington, United States

[c]University of Pittsburgh, Department of Radiology, Pittsburgh, Pennsylvania, United States

[d]University of Pennsylvania, Department of Radiology, Philadelphia, Pennsylvania, United States

A theory [or simulation] is something nobody believes, except the person who developed it. An experiment is something everybody believes, except the person who performed it.

-Albert Einstein

Essentially, all models are wrong, but some are useful.

-George Box

Why Do We Need VCT?

In all scientific disciplines, experimentation is a hallmark of scientific rigor, where theories, assumptions, and derivations can be verified and refined by data extracted from designed interrogation of physical reality. In medicine, however, that possibility is highly constrained, for many good reasons. Researchers cannot simply "experiment" on humans. Certain types of experiments are allowed, provided they adhere to strong ethical and regulatory standards after a rigorous review processes such as those provided by institutional review boards. However, because of those very appropriate standards, many experiments are simply undoable. Some investigations have thus been delegated to the animal domain – through animal studies which are also constrained by necessary ethical limitations of their own. And still, for human studies or clinical trials that are approved, there are often sharp constraints imposed by the limits of the time allowed for accrual, available personnel, extensive paperwork, and financial resources. This is due to the high heterogeneity among human subjects, where each person presents unique individual factors that can confound the results. Clinical trials typically aim to minimize the confounding effects by strict inclusion-exclusion criteria that organizes the subjects into delimited cohorts, but not all confounding effects can be avoided. Consequently, a trial needs to accrue a large enough number of subjects to overcome the confounding affects and achieve statistical significance across variabilities. This results in a significant cost that is ultimately born by the public, whether through tax dollars (e.g., National Institutes of Health), philanthropy (e.g, foundations), or industry which eventually channels those costs to the customers.
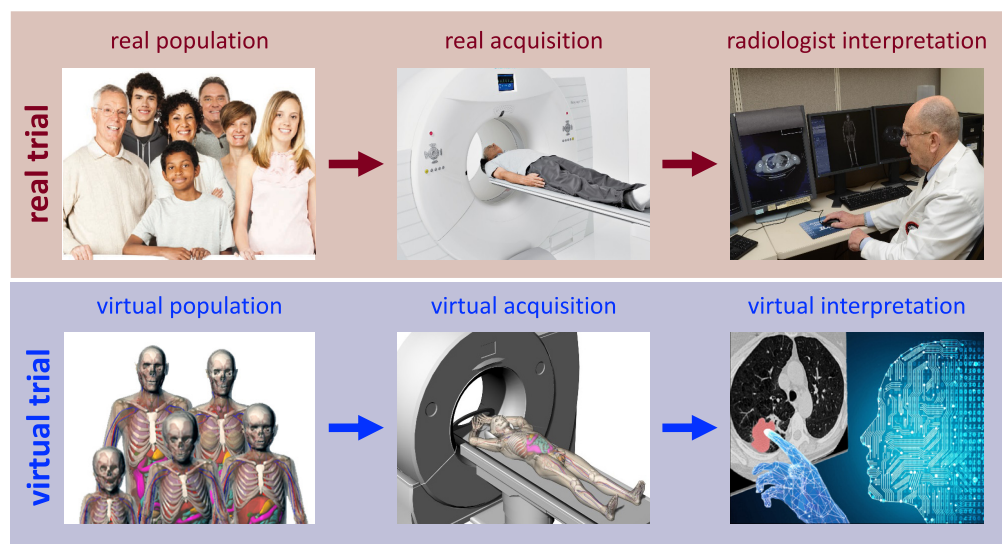
When it comes to medical imaging, the limitations of clinical trials take additional dimensions, figuratively and literally. In imaging, the development and use of the technologies tend to be highly complex. For example, the development of a new CT system involves multiple considerations that influence the balance between resolution, noise, and quantitative accuracy, considerations such as regularization factors in iterative reconstructions or energy thresholds in photon-counting acquisitions. Likewise in clinical use, the said system presents a wide array of selectable features and parameters such as the choice of kV, reconstruction kernel, dose level, or flux modulation. Image quality and effectiveness is governed by these factors, which should ideally be selected for optimum care and safety of the patient, recalling that quality healthcare should always take place in the context of the *individual* patient. However, to this day, while this aspiration has been approached, it has remained largely unfulfilled. This is because of the fact

that, even if a trial is deemed ethically feasible, the myriad of user-defined parameters are simply too numerous to be systematically or even approximately investigated in a clinical trial without making the trial prohibitively unwieldy. And in the development of the technology, the prototype systems are simply too unstable for human trials, even with additional levels of compliance, such as the investigational device exception process invoked by the FDA. Consequently, the development or use choices in imaging are often made based on assumptions (ideally educated ones), predictions from simple phantoms (often crude predictions based on clinically unrealistic phantoms), or concordance with peer practice. These are all reasonable starting points but no guarantors of optimality.

This challenge of undoability of human trials for medical imaging is further compounded by the multiplicity of clinical tasks. Each clinical task requires a different combination of imaging factors for optimal performance. The optimum imaging protocol for liver lesions is very different from that for pulmonary embolism. In imaging technology, system specifications are often taken as indicators of clinical quality, e.g., a faster system or one with smaller pixel sizes provide higher quality data. But what is considered better in terms of specifications does not necessarily translate to effectiveness for a clinical task or across the diversity of clinical tasks. And further adding to this challenge is the emerging demand in imaging to not only ascertain *interpretive* tasks, but also to provide robust and precise *quantitative* information as biomarkers for therapeutic management and improved outcomes. How to determine the precise design and use choices to obtain optimum and consistent task-based imaging output across the diversity of patients, interpretive and quantitative tasks, and imaging systems has remained largely uncertain.

What is VCT?

Within the context of the aforementioned needs and challenges, virtual imaging clinical trial (VCT) is an emerging methodological alternative to clinical trial in imaging. A VCT is essentially an analogue to a clinical trial in that the population of human subjects is replaced with a population of virtual digital subjects, the imaging systems with virtual imaging simulators, and the clinical interpretations with virtual interpretations. As illustrated in Fig. 1, in VCT, the virtual "subject" is then virtually "imaged" and the image is virtually "interpreted," emulating the clinical process without an actual clinical trial. A VCT offers a feasible and efficient means to conduct experimentation in medical imaging without the associated ethical, time, and cost limitations. As such, a VCT provides the practical ability to assess and optimize a host of design factors in the development and use of imaging technologies, a task not possible through clinical trials.

Courtesy of Ravin Labs, Duke University

Fig. 1 A clinical trial process (top) emulated by a virtual clinical trial (bottom).
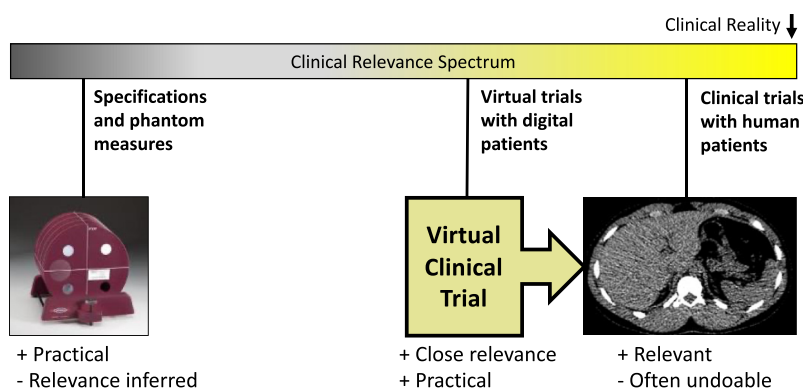
Fig. 2 Virtual clinical trials in imaging (middle) offer an intermediary alternative to simpler but simplistic phantom-based models (left) and more accurate but resource-intensive clinical trials (right).

Methodologically, any form of experimentation in medicine, including a clinical trial, can be thought as an approximation of clinical reality. System specifications and characterizations through non-anthropomorphic phantoms also provide approximations, but approximations that are far removed from clinical reality. Within this spectrum of approximations from specifications and phantom measures to clinical trials, as illustrated in Fig. 2, a VCT can be understood as yet another form of approximation. Within this mental map, a VCT can be thought as an interpolation point, providing a representation of clinical reality that is much more realistic than what is possible through simpler surrogates of geometrical phantoms and measurements, while at the same time circumventing the major limitations of clinical trials. In that way, a VCT can be used as a practical and reliable alternative to a clinical trial, either as a reflection of the clinical reality on its own, or as a precursor to enable better targeted subsequent clinical trials, if such trials are still deemed feasible and needed.

While time-cost efficiency and ethical feasibility are the main advantages of VCTs, VCTs offer one additional attribute that even exceeds what is possible through actual clinical trials: the perfect knowledge and the precise control over the *ground truth*. As the condition of the patient is defined *a priori*, a VCT makes it possible to ascertain how an image rendition represents the ground truth. This is a unique capability that can never be assured perfectly in clinical trials. For most imaging trials, the ground truth of the patient is unknown. For example, when imaging a lesion, the exact morphology of the lesion in unknown. Even if the trial includes surgical assessment of the lesion, the process of surgical resection would alter the resected lesion – thus the *actual* morphology of the lesion, present at the time of imaging, can *never* be known precisely. If one is to determine the suitability of an image hardware or software to ascertain the volume of a tumor, the precise volume *in vivo* cannot be known. But the precise lesion volume in VCT is known *in silico*. Thus one can more precisely ascertain the utility of the said hardware or software.

For the digital ground truth of the VCT to be trusted as a relevant ground truth reflective of the actual reality, it is of critical importance to ensure that the VCT closely models that reality. Otherwise, a VCT cannot be assumed to offer much advantage over simpler phantom investigations (in Fig. 2 illustration, creating a methodology close to the right side of the spectrum). Each component of a VCT should strive to realistically reflect its corresponding physical analogue: The digital subject needs to adequately represent the anatomical and physiological reality of the individual patient as well as the diversity across patients; the imaging simulator needs to echo the performance of the modeled imaging system and the variations across different systems; and the virtual interpreter needs to emulate the discernibility of the anticipated clinical interpreter, either the human observer or observers in the case of visual interpretation, or the computational "observer" in the case of image analysis algorithms.

Recent years have witnessed a steady progression in enhancing the realism of human models, imaging models, and observer models. So much so that in fact their integration has become a natural progression, promoting the emergence of the VCT as a distinct methodological system of its own. This coming to maturation of the VCT as a field is further demonstrated by multiple

publications, short courses on in different medical imaging subfields (in mammography, CT, and PET), and even multiple FDA approvals based on VCT studies.

In this Issue

This special section of Volume 7 Issue 4 of the *Journal of Medical Imaging* highlights multiple projects that develop and use virtual clinical trials to answer relevant clinical questions.

Abadi *et al.* in "Virtual clinical trials in medical imaging: a review," authored by senior investigators from multiple institutions, offer a systematic review of the VCT methodologies and applications across diverse modalities of X-ray, nuclear, MRI, and ultrasound imaging. The review includes types and aspects of simulations for each modality, formation of digital phantoms, and the evaluation of the resulting images. It further highlights example VCTs, methods of validations, and a perspective on future direction of VCTs.

Mattonen *et al.* in "Quantitative imaging feature pipeline: a web-based tool for utilizing, sharing, and building image-processing pipelines" detail a software resource facilitating processing uploaded image data or programs. This resource can facilitate the analysis and management of VCT images, processes, and algorithms.

Abadi *et al.* in "Virtual clinical trial for quantifying the effects of beam collimation and pitch on image quality in computed tomography" apply the VCT methodology to CT, to answer the question of how increased beam collimation and pitch, which speed the scanning speed and reduce motion artifact, might negatively influence the spatial rendition of lung lesions. The subject was imaged repeatedly across a range of respiratory and cardiac motion, something that is possible only through VCTs.

Gong *et al.* in "Deep-learning-based model observer for a lung detection task in computed tomography" devise and evaluate a virtual interpretation model, an observer model based on machine learning. The model was applied across a range of imaging condition for the detection of lung lesions in CT, demonstrating that the model can closely emulate the results from human radiologists.

Zeng *et al.* in "Computational reader design and statistical performance evaluation of an in-silico imaging clinical trial comparing digital breast tomosynthesis with full-field digital mammography" apply VCT in the context of breast imaging, ascertaining the detection of masses and macrocalcifications by breast tomosynthesis and mammography, across 3000 virtual breasts, echoing the prior clinical indications that DBT has higher performance in mass detection and affirming the validity of channelized observer models for like breast imaging tasks.

Finally, Vancoillie *et al.* in "Verification of the accuracy of a hybrid breast imaging simulation framework for virtual clinical trial applications" tackle the same modalities of the paper by Zeng *et al.* but focuses on creating realistic lesion models in existing 2D and 3D breast images. Their results suggest that the modeled lesions and combined lesion-breast constructs can realistically model DBT and mammography imaging scenarios for VCT

What is Next?

Beyond the focused topics in this issue, VCT research has a wide range of applications, across all imaging modalities. Those include ascertaining and mitigating variability in the biological, patient, acquisition, image reconstruction, processing, and interpretation processes; optimizing existing imaging methods; conducting virtual trials across the range of specific clinical needs; and evaluating new imaging hardware and software.

Research into and with VCTs still has a long developmental path ahead. New developments are anticipated and encouraged (by us) across a range of needs such as

- improving the realism of digital phantoms,
- microscopic biological modelling, both anatomical and functional,
- developing most relevant metrics for VCT reporting,
- developing commons standards for VCT processes and reporting,

- developing pipelines for data handling and processing,
- enhancing computation efficiency,
- validation methods against physical or clinical measurements,
- ascertaining the degree of closeness to physical or clinical reality to ensure relevance,
- ascertaining the degree of closeness needed for maximum clinical relevance, and
- ascertaining the generalizability of VCT findings.

We encourage the readers of this special section to ponder and tackle these opportunities and potentials. We are excited about the notable advantages that VCTs can offer our medical research and healthcare through its realism and pragmatism. Enhanced utility and advancement in the VCT methodology is distanced to lead to a measurable impact on the reach of medical imaging, to enrich medical research, and to advance human health.

Acknowledgments

Ehsan Samei, PhD, DABR, FAAPM, FSPIE, FAIMBE, FIOMP is an imaging scientist and professor of radiology, physics, medical physics, BME, and ECE at Duke University, where he also serves as the chief imaging physicist for the Duke University Health System. His interests include virtual clinical trials and clinically relevant metrology of imaging quality and safety for optimum interpretive and quantitative performance. He strives to bridge the gap between scientific scholarship and clinical practice through (1) meaningful realization of translational research and (2) actualization of clinical processes that are informed by scientific evidence.

Paul E. Kinahan, PhD, FIEE, FAAPM, FSNMMI is the vice-chair for research and a professor of radiology and bioengineering at the University of Washington. He is also the clinical director of PET/CT imaging physics for UW Medicine. His research includes optimizing the physics of PET and CT imaging, the use of statistical image reconstruction methods, objective assessment of image quality, and the use of quantitative analysis in imaging-based clinical trials.

Robert M. Nishikawa, PhD, FAAPM, FSBI, FAIMBE, FSPIE is a medical physicist and professor of radiology at the University of Pittsburgh. His research focuses on quantitative imaging, including computer-aided diagnosis, and technology assessment, both principally in breast imaging.

Andrew D. A. Maidment, PhD, DABR, FAAPM is an associate professor of radiology at University of Pennsylvania, and chief of physics at University of Pennsylvania Health System. Maidment has focused his research on developing novel x-ray imaging systems for more than 30 years, and began studying virtual clinical imaging trials as a way of designing imaging systems optimized to specific imaging tasks 21 years ago with his close colleague Predrag Bakic. Today, their OpenVCT simulation platform is used world-wide.