Commentary

For reprint orders, please contact: reprints@futuremedicine.com

Personalized Medicine

How artificial intelligence could redefine clinical trials in cardiovascular medicine: lessons learned from oncology

Chayakrit Krittanawong*[,1], Kipp W Johnson[2] & WH Wilson Tang[3,4,5]

[1]Department of Internal Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA

[2]Department of Genetics & Genomic Sciences, Institute for Next Generation Healthcare, Icahn School of Medicine at Mount Sinai, New York, NY, 10029, USA

[3]Department of Cardiovascular Medicine, Heart & Vascular Institute, Cleveland Clinic, OH, 44195, USA

[4]Department of Cellular & Molecular Medicine, Lerner Research Institute, Cleveland, OH, 44195, USA

[5]Center for Clinical Genomics, Cleveland Clinic, Cleveland, OH, 44195, USA

*Author for correspondence: Tel.: +1 212 523 4000; Fax: +1 212 523 8605; Chayakrit.Krittanawong@mountsinai.org

First draft submitted: 1 November 2018; Accepted for publication: 10 December 2018; Published online: 6 March 2019

Keywords: artificial intelligence • clinical trials • deep learning • machine learning • precision cardiovascular medicine • precision medicine

Features & limitations in randomized controlled trials

The randomized controlled trial (RCT) is the gold standard for evidence in clinical research. RCTs allow an investigator to test a specific hypothesis by controlling for known confounders via regression analysis and ameliorating the effect of potential unknown confounders via randomization. Theoretically, RCTs provide unbiased estimates of a relative treatment effect when conducted with adequate number of subjects [1]. Ideally, each and every medical intervention would be subjected to rigorous and well-powered RCTs to demonstrate effectiveness before the interventions are implemented into clinical practice. However, practically, this ideal is unattainable for several reasons, the most important of which are the immense expense of planning and executing an RCT; difficulty recruiting patients to participate and challenges in generalizing findings to those that have been excluded from the RCT. In fact, today, over 30,000 cardiovascular disease (CVD) related clinical trials in clinicaltrials.gov have either not been completed or have not reported results in peer-reviewed publications because of limited power (i.e., difficulty in recruiting potential candidates), inadequate period of follow-up, heterogeneous study populations with unaccounted inter-individual variabilities, adverse events that were potentially unrelated to interventions or null publication bias. These deficiencies often lead to difficulties interpreting trial results or generalizing trial findings to individual patients at the point of care [2,3]. Furthermore, many RCTs deviate from the ideal analytic strategies because of design flaws. Importantly, clinical trials often enroll patients who may not be representative of the targeted population in the first place [4]. RCTs also cannot reveal the distribution of risk in the overall population, nor provide accurate measures of absolute effectiveness for a treatment when the treatment effects can differ between the clinical trial and the real world (e.g., medication adherence, observer effects, patient characteristics or drug interactions). For these reasons, there has been an ongoing search for better experimental approaches.

In this paper, we will discuss the potential for adaptive, artificial intelligence (AI) guided clinical trials to generate evidence in cardiovascular medicine. AI may allow an investigator to predict results sooner and potentially harming fewer patients, or to perform AI-assisted randomization strategies using a high-dimensional set of variables.

Beyond complex adaptive clinical trials: from oncology to cardiovascular medicine

Much existing work has focused upon improving conventional clinical trials by making them more flexible and efficient. These strategies typically involve incorporating findings that occur during the trial to modify or refine the trial design. For example, the concept of an adaptive trial allows modifications to the trial or statistical procedures discovered during an interim analysis [5]. Adaptive trials can be classified into several categories based upon their specific mechanisms: sample size/baseline characteristic/outcome reassessments; treatment arm modification; hypothesis/design modification based upon interim data and prospective–retrospective or enrichment designs.

In addition, the statistical concept of adaptive trials is usually not a statistically significant p-value, but interim data-driven modification with methodologies such as Markov chain Monte Carlo, decision rules or Bayesian methodologies. To date, adaptive trials (e.g., BATTLE-2 [6], FOCUS4 trial [7], I-SPY trial [8], NCI-MATCH [9], Lung-MAP SWOG S1400 [10] and DIAN-TU [11]), N-of-1 trials (single-patient trial designs), basket trials and umbrella trials (which operate on subsets of patients) have been completed or are ongoing. However, the methods of these innovative trial designs have not yet been widely applied or considered in cardiovascular medicine [12]. With the increasing utilization of genomic-based diagnosis in the heterogeneous disease of cancer, oncologists have begun to more widely utilize N-of-1 trials to treat their patients. Oncological clinical trials such as study B2225 using Imatinib for ten patients (eight with locally advanced dermatofibrosarcoma protuberans and two with metastatic disease) to assess the activity of imatinib in diseases associated with expression or activation of imatinib-sensitive tyrosine kinases demonstrated the potential of this approach in rare disease [13]. We believe that similar approaches may be of value in cardiovascular medicine, such as in the example of a trial of a new heart failure medication, for in peripartum cardiomyopathy.

Interestingly, CVDs are equally heterogeneous even before consideration of patient genetics results, and arguably even more heterogeneous when incorporating genetic results or biological features as defined by biomarkers. However, current cardiovascular treatments still poorly, they incorporate patient stratification and often resort to the 'one-size-fits-all' doctrine. In the case of heart failure, recent 'phenomapping' efforts have highlighted the potential for redefining subtypes of heart failure with preserved ejection fraction to different clusters that may differentially respond to different treatments [14,15]. Meanwhile, reduced left ventricular ejection fraction (<35–40%) has been the key inclusion criterion that identified therapeutic responses, yet factors identified in *post-hoc* analyses like ventricular sizes, rhythm or conduction patterns or valvular functions may influence the response rates of intervention that are often not accounted for in patient selection or treatment adjustments. Moreover, to date most genetic testing in CVD occurs only in specific populations such as those at high risk of sudden death or possessing suspected inherited CVD. As genome-based cardiovascular medicine becomes more widespread, the potential for incorporating genetics into cardiovascular trial design will increase dramatically. Finally, we note that there are methodological challenges for adaptive clinical trials that need to be addressed (Supplementary Table 1).

Types of artificial intelligence: machine learning, deep learning

Briefly, machine learning (ML) can be classified into supervised (i.e., classification) and unsupervised learning (i.e., cluster analysis) [16,17]. Deep learning (DL) is an especially promising subfield of ML that has developed more recently with the advent of wide accessibility to highly parallel computing power via computer graphics processing units. Although ML may be used for classification or to cluster patients by many variables simultaneously, as of now there exists no evidence or examples to demonstrate that AI is better than traditional statistical methods in clinical medicine. The concept of biomarker-driven adaptive trials such as BATTLE-2 trial [6] could be categorized into training, testing and validation phase which are similar to the practices of ML. Therefore, advances in AI such as DL may ultimately prove beneficial when applied to adaptive trials. Indeed, there are many examples of ML and DL already in use in medicine [18].

Meanwhile, deep reinforcement learning (DRL), an especially promising but still emerging field of machine learning which is often used in computer gaming could also be applied to RCT methodologies. DRL could potentially overcome ethical limitation of current RCT (e.g., different in outcomes between two arms) because DRL can learn from prior experience by selecting actions and continue to select better actions to improve outcomes. For example, DRL can help agents (such as a clinical making decision tool) to select more accurate actions to reduce readmission rate in heart failure patients by analyzing current results to predict the future outcomes, and continuously refining such algorithms.

Application of AI in cardiovascular clinical trials

The prospect of future trials using AI-guided biomarkers or genomics in advanced adaptive trials may impact all components of the clinical trial process, including: study design; study recruitment; study conduct and study intervention and interpretation. Supplementary Table 2 enumerates the potential of AI in different clinical trial methodologies based on types of clinical trials.

Study design

Study design in clinical trial is a challenging subject. The ability to modify ongoing trials using interim data in adaptive designs is promising but more robust statistical analysis without bias is needed to use predictive analytics. Study design is crucial for clinical trials to determine the accuracy, reliability, power and generalizability. The main advantage of adaptive trials is to simulate multiple rounds of analyses to make sure the benefits of the design substantially outweigh the potential risks. Thus, with powerful predictive ability, AI could be used to simulate a virtual round of data analysis (i.e., tenfold validation or the parametric G formula) [19]. AI will play an important role in the statistical planning (i.e., the predictive analytics in interim analysis and the final analysis). AI could potentially be used to operationalize (i.e., through identification of stratified cohorts) or optimize study design (i.e., multiple biomarkers, multiple treatments and multiple parameter from different modalities). The minimization of uncertainty when optimizing study design is one possible target for improvement. However, at present we believe that the main limitation of current clinical trial is generalizability and external validity. Improved methodologies such as Markov chain Monte Carlo, decision rules, Bayesian analysis strategies or future novel AI techniques may be useful in study design. AI could help with identifying targeted populations by selecting multiple variables (i.e., different ethnicity, life style, diet, culture, income, geographical location, number of emergency department or primary care doctor visits) using sophisticated techniques and ultimately improve the external validity of the study's findings. AI could potentially identify drug efficacy, identify adverse reactions or reduce both the time and the costs compared with more traditional clinical trials. For example, applying methods drawn from AI to wearable technology has the potential to incorporate real-world variables into protocol designs. Most importantly, AI can facilitate genome driven basket/umbrella clinical trials to select genes/single nucleotide variants in small population with rare heterogeneous syndromes to enhance power and statistical methods.

Study recruitment

At present, the highest frequency application of AI in clinical trials is for the recruitment of patients to studies. For example, there are many commercial applications that already claim to use AI to some extent for effective study recruitment. However, this is a fertile area and there remain many possibilities offered by AI-recruited trials be explored. The AI company Deep 6, for example, uses natural language processing to read physicians' notes, pathology reports, diagnoses, recommendations and to detect hard-to-find lifestyle data, such as smoking and activity history to ultimately match patients to clinical trials with appropriate criteria. Mendel.ai, another emerging AI platform, uses an natural language processing algorithm that interprets unstructured data pulled from clinicaltrials.gov to match qualified patients with relevant clinical research opportunities, ultimately resulting in a reduction in time and cost [20]. AI could potentially be used to identify eligible patients using inclusion and exclusion criteria from electronic health records (EHRs) [21,22] or define inclusion and exclusion criteria based on results of similar trials in the past. Although commercial applications from IBM such as Watson for clinical trial matching claim the ability to extract targeted populations for clinical trials [23], studies have shown that there are biases in AI system including IBM Watson's predictive algorithm in a substantial number of cases based on predictive analytics [24–27]. This was because the methodologies often times did not incorporate standard evaluation techniques, and thus there is no way to validate their rigor and/or reproducibility. AI-platforms such as these could potentially be used to overcome limitations of previous clinical trials. Replication of the SPRINT trial [28], for example, could be used with AI to recruit candidates, titrate and optimize their BP [19]. For example, a known limitation of SPRINT is that more than 90% of the trial's patients were under antihypertensive drug treatment at baseline and this baseline BP with antihypertensive drug treatment could confound results. Thus, AI can be used to stratify baseline BP using real-time measurement from wearable technology that may identify latent variables (i.e., medication compliance, average time of BP) and this method could potentially better stratify groups of patients in the SPRINT trial.

Study conduct

The utilization of AI to ensure rigor and reproducibility of results is promising. AI could potentially facilitate changing how RCTs are conducted. Concepts such as randomization, double-blinding and control group comparisons are the strongest features of RCTs. AI could help to generalize the results. For example, AI can analyze EHR or biobanks from drugs sensitivity to predict dropout rate ahead-of-time, which could then be incorporated into aggressive education and intervention in a group of high-risk patient in terms of dropout rate [29]. AI can be used to perform exploratory analysis from vital signs, healthcare providers' notes in EHR (i.e., physicians, consulting services, social worker, physical therapy, and register nurses), biomarkers, omics data (i.e., genomics, transcrip-

tomics, metabolomics, proteomics, more recently 'exposomics') and cardiac imaging, leading to novel data driven hypotheses, new discovery of phenotypes or pathogenesis of CVDs. In addition, AI may help in terms of direct observation or supervision during trial conduct. A combination of wearable technology for real-time variables and cloud-based system to conduct study has a potential to solve current adaptive trials problems. Finally, AI is used in drug discovery. A large number of commercial startups have already conducted clinical trials using drugs discovered via AI methods [30].

Study intervention & interpretation

AI-directed clinical trial may perform better than traditional clinical trial because AI often has increased predictive analytic ability compared with humans. In general, interventions in RCTs are *a priori* predicted to have beneficial effect. AI could be used to more rapidly identify negative results in ongoing RCTs. Predictive analytics is key for AI-guided study intervention before the trial and during interim analysis. Ongoing trials using AI in medicine such as SYNERGY-AI trial (NCT03452774) are underway. Another potential is to use AI to shorten time of trials while maximizing benefits. For example, in the COAPT trial [31], AI could potentially predict HF hospitalization in guideline-directed medical therapy and MitraClip at 12 months rather than 24 months. In addition, AI could help with process automation (i.e., tailor online surveys, automated inform consent, reduce duplicate data entry, optimize dose of medication and drop ineffective treatments) [32,33].

Future perspective

Overall, with emerging wearable technology and big data, AI has potential to enable innovative trials a faster development of new interventions. In particular, the incorporation of wearable device technology may be helpful in novel designs to study new drugs by capturing unique data such as better monitoring their potential side effects or tracking real-time medication adherence [34]. AI-guided wearable technology could be used to titrate medication using real-time vital signs or during interim analysis to more precisely assess real-time medication adherence. The advent of upcoming trials using secure wearable technology and storing data in shareable cloud systems or block chain, a huge, public, secure and decentralized datastore, is promising [35]. In addition, AI-guided clinical decision support using multi-omics approaches to maximize effective treatment group or minimize ineffective treatment group during interim analysis. Finally, AI-guided clinical trials using multiple imputation, cross validation or data augmentation techniques could be used to explore unknown pattern of rare disease such as premature atherosclerosis, spontaneous coronary artery dissection or arrhythmogenic right ventricular cardiomyopathy. To conclude, we believe that AI could and will improve all aspects of the modern clinical trial by incorporating more sophisticated modeling using the depth and breadth of data that are beyond traditional statistical methods.

Conclusion

In conclusion, although AI has a potential to revolutionize RCT in several areas, the implementation of AI in RCT much remains to be explored. The attempt to apply AI in RCTs is promising and could increase the efficacy of RCTs in cardiovascular medicine. The practical guide using AI-guided clinical trials for future investigators is needed to be established.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/full/10.2217/pme-2018-0130

Acknowledgments

The authors thank Martin S Tallman (Leukemia Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, NY, USA), for oncology consultations and criticism of the manuscript.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Sedgwick P. What is a non-inferiority trial? *BMJ (Online)* 347 (2013). www.bmj.com/content/347/bmj.f6853
2. Hwang TJ, Carpenter D, Lauffenburger JC, Wang B, Franklin JM, Kesselheim AS. Failure of investigational drugs in late-stage clinical development and publication of trial results. *JAMA Intern. Med.* 176(12), 1826–1833 (2016).

•• **Provides an example of limitations in current clinical trials.**

3. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to translate into benefits for patients. *Trials* 18(1), 122 (2017).
4. Bhatt DL, Mehta C. Adaptive designs for clinical trials. *N. Engl. J. Med.* 375(1), 65–74 (2016).

•• **A useful review about adaptive clinical trials.**

5. Chow SC, Chang M, Pong A. Statistical consideration of adaptive methods in clinical development. *J. Biopharm. Stat.* 15(4), 575–591 (2005).
6. Papadimitrakopoulou V, Lee JJ, Wistuba II *et al.* The BATTLE-2 study: a biomarker-integrated targeted therapy study in previously treated patients with advanced non-small-cell lung cancer. *J. Clin. Oncol.* 34(30), 3638–3647 (2016).
7. Kaplan R, Maughan T, Crook A *et al.* Evaluating many treatments and biomarkers in oncology: a new design. *J. Clin. Oncol.* 31(36), 4562–4568 (2013).
8. Barker AD, Sigman CC, Kelloff GJ, Hylton NM, Berry DA, Esserman LJ. I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clin. Pharmacol. Ther.* 86(1), 97–100 (2009).
9. Barroilhet L, Matulonis U. The NCI-MATCH trial and precision medicine in gynecologic cancers. *Gynecol. Oncol.* 148(3), 585–590 (2018).
10. Herbst RS, Gandara DR, Hirsch FR *et al.* Lung master protocol (Lung-MAP) – a biomarker-driven protocol for accelerating development of therapies for squamous cell lung cancer: SWOG S1400. *Clin. Cancer Res.* 21(7), 1514–1524 (2015).
11. Bateman RJ, Benzinger TL, Berry S *et al.* The DIAN-TU next generation Alzheimer's prevention trial: adaptive design and disease progression model. *Alzheimers Dement.* 13(1), 8–19 (2017).
12. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Per. Med.* 8(2), 161–173 (2011).
13. Heinrich MC, Joensuu H, Demetri GD *et al.* Phase II, open-label study evaluating the activity of imatinib in treating life-threatening malignancies known to be associated with imatinib-sensitive tyrosine kinases. *Clin. Cancer Res.* 14(9), 2717–2725 (2008).
14. Shah SJ, Katz DH, Selvaraj S *et al.* Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* 131(3), 269–279 (2015).
15. Krittanawong C, Kukin ML. Current management and future directions of heart failure with preserved ejection fraction: a contemporary review. *Curr. Treat Options Cardiovasc. Med.* 20(4), 28 (2018).
16. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial intelligence in precision cardiovascular medicine. *J. Am. Coll. Cardiol* 69(21), 2657–2664 (2017).

•• **A useful review about artificial intelligence in cardiovascular medicine.**

17. Johnson KW, Soto JT, Glicksberg BS *et al.* Artificial intelligence in cardiology. *J. Am. Coll. Cardiol* 71(23), 2668–2679 (2018).
18. Krittanawong C, Bomback AS, Baber U, Bangalore S, Messerli FH, Tang WW. Future direction for using artificial intelligence to predict and manage hypertension. *Curr. Hypertens. Reps.* 20(9), 75 (2018).
19. Johnson KW, Glicksberg BS, Hodos RA, Shameer K, Dudley JT. Causal inference on electronic health records to assess blood pressure treatment targets: an application of the parametric g formula. *Pac. Symp. Biocomput.* 23, 180–191 (2018).
20. Buhr S. Mendel.Ai nabs $2 million to match cancer patients with the latest clinical trials. (2017). https://techcrunch.com/2017/07/01/mendel-ai-nabs-2-million-to-match-cancer-patients-with-the-latest-clinical-trials/
21. Miotto R, Weng C. Case-based reasoning using electronic health records efficiently identifies eligible patients for clinical trials. *J. Am. Med. Inform. Assoc.* 22(e1), e141–150 (2015).
22. Glicksberg BS, Miotto R, Johnson KW *et al.* Automated disease cohort selection using word embeddings from electronic health records. *Pac. Symp. Biocomput.* 23, 145–156 (2018).
23. Fassbender M. Outsourcing-Pharma. AI drives 80% increase in clinical trial enrollment. (2018). www.outsourcing-pharma.com/Article/2018/03/26/AI-drives-80-increase-in-clinical-trial-enrollment
24. CRaI S. Ibm's watson supercomputer recommended 'unsafe and incorrect' cancer treatments, internal documents show. (2018). www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/
25. Buolamwini J, Gebru T. Gender shades: intersectional accuracy disparities in commercial gender classification. Presented at: *Conference on Fairness, Accountability and Transparency.* NY, USA. 77–91 (February 23 and 24th, 2018).
26. Madaan N, Mehta S, Agrawaal T *et al.* Analyze, detect and remove gender stereotyping from bollywood movies. Presented at: *Conference on Fairness, Accountability and Transparency.* NY, USA. 92–105 (February 23 and 24th, 2018).

27. Srivastava B, Rossi F. Towards composable bias rating of AI services. Presented at: *AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society.* LA, USA (February 2-3, 2018).

28. Wright JT Jr., Williamson JD, Whelton PK *et al.* A randomized trial of intensive versus standard blood-pressure control. *N. Engl. J. Med.* 373(22), 2103–2116 (2015).

29. Glicksberg BS, Johnson KW, Dudley JT. The next generation of precision medicine: observational studies, electronic health records, biobanks and continuous monitoring. *Hum. Mol. Genet.* 27(R1), R56–R62 (2018).

30. S S.117startups using artificial intelligence in drug discovery. (2018). https://blog.benchsci.com/startups-using-artificial-intelligence-in-drug-discovery

31. Mack MJ, Abraham WT, Lindenfeld J *et al.* Cardiovascular outcomes assessment of the mitraclip in patients with heart failure and secondary mitral regurgitation: design and rationale of the COAPT trial. *Am. Heart J.* 205, 1–11 (2018).

32. Bain EE, Shafner L. Use of a novel artificial intelligence platform on mobile devices to assess dosing compliance in a Phase II clinical trial in subjects with schizophrenia. 5(2), e18 (2017).

33. Shiovitz TM, Bain EE, McCann DJ *et al.* Mitigating the effects of nonadherence in clinical trials. *J. Clin. Pharmacol.* 56(9), 1151–1164 (2016).

34. Krittanawong C, Johnson KW, Hershman SG, Tang WHW. Big data, artificial intelligence, and cardiovascular precision medicine. *Expert Rev. Preci. Med. Drug Dev.* 3(5), 305–317 (2018).

- **Provides an example of a potential big data in cardiovascular medicine.**

35. Peters GW, Panayi E. Understanding modern banking ledgers through blockchain technologies: future of transaction processing and smart contracts on the internet of money. In: Paolo Tasca, Tomaso Aste, Loriana Pelizzon, Nicolas Perony (Eds). *Banking Beyond Banks and Money.*SpringerNature, Switzerland AG 239–278 (2016). https://link.springer.com/book/10.1007/978-3-319-42448-4