Computational Medicine, Present and the Future: Obstetrics and Gynecology Perspective

Radek Bukowski MD, PhD; Karl Schulz, PhD; Kelly Gaither, PhD; Keri K. Stephens, PhD; Dave Semeraro, PhD; Justin Drake, PhD; Gordon Smith, MD, PhD; Craig Cordola, FACHE; Thaleia Zariphopoulou, PhD; Thomas J.R. Hughes, PhD; Christopher Zarins, MD; Dimitri Kusnezov, PhD; Rep. Donna Howard, M.A.; Tinsley Oden, PhD

Department of Women's Health, Dell Medical School, University of Texas at Austin (Radek Bukowski, Karl Schulz, Kelly Gaither);

Oden Institute for Computational Engineering and Sciences (Karl Schulz, Thomas J.R. Hughes, Tinsley Oden)

Texas Advanced Computing Center (Kelly Gaither, Dave Semeraro, Justin Drake)

Moody College of Communication, The University of Texas at Austin (Keri K. Stephens)

Department of Obstetrics and Gynaecology, University of Cambridge, UK (Gordon Smith)

Executive Vice President and Chief Operating Officer Ascension (Craig Cordola)

Department of Mathematics and IROM, McCombs School of Business at The University of Texas at Austin (Thaleia Zariphopoulou)

HeartFlow, Inc (Christopher Zarins)

Deputy Under Secretary for Artificial Intelligence & Technology US Department of Energy (Dimitri Kusnezov)

Texas House of Representatives, Representative the 48th District (Rep. Donna Howard)

Corresponding Author: Radek Bukowski MD, PhD, Department of Women's Health, Dell Medical School, University of Texas at Austin radek.bukowski@austin.utexas.edu

Disclosures: The authors report no conflict of interest.

Short title: Computational Medicine
Word Count: Abstract 286 **Text** 11,398

ABSTRACT

Medicine is, in its essence, decision making under uncertainty; The decisions are made about tests to be performed and treatments to be administered. Traditionally the uncertainty in decision making was handled using expertise collected by individual providers, and more recently systematic appraisal of research in the form of evidence-based medicine. The traditional approach has been successfully used in medicine for a very long time. However, it has significant limitations due to the complexity of the system of the human body and health care. The complex systems are networks of highly coupled components intensely interacting with each other. These interactions give those systems redundancy, thus robustness to failure, and, at the same time, equifinality, that is, many different causative pathways leading to the same outcome. The equifinality of the complex systems of the human body and health care system demand the individualization of medical care, medicine, and medical decision making. Computational models excel in modeling complex systems and, in consequence, enabling individualization of medical decision making and medicine. Computational models are theory- or knowledge-based models, data-driven models, or models that combine both approaches. Data are essential, although to a different degree, for computational models to successfully represent complex systems. The individualized decision making, made possible by the computational modeling of complex systems, has the potential to revolutionize the entire spectrum of medicine from individual patient care to policymaking. This approach allows applying tests and treatments to individuals who receive a net benefit from them, for whom benefits outweigh the risk, rather than treating all individuals in a population because on average the population benefits. Thus, the computational-modeling-enabled individualization of medical decision making has the potential to both improve health outcomes and decrease the costs of health care.

KEY WORDS: computation, modeling, data, uncertainty, machine learning, data-driven models, theory-based models, physics-based models,

CONDENSATION: Computational-modeling-enabled individualization of medical decision making has the potential to both improve health outcomes and decrease the costs of health care.

"*Medicine is a science of uncertainty and an art of probability.*"[1] This is how Sir William Osler described medicine over a century ago, a statement that resonates with us as strongly today. The considerable and ever-present multi-faceted uncertainty, in comparison to other disciplines, is what makes the practice of medicine so challenging. This is mainly because in medicine we face a very complex system and we lack fundamental natural laws describing its function to inform our models. The uncertainty and complexity are especially pronounced in the practice of obstetrics, where two or more young and usually otherwise healthy patients with long life expectancies and with frequently conflicting health interests are being cared for and where the stakes and expectations are very high[2, 3]. Thus, medicine is, in its essence, decision making under uncertainty; decisions about the tests to be performed and treatments to be administered. The inescapable and pervasive uncertainty is the consequence of two issues facing every, although to a different degree, medical decision: the incomplete information to make the decision and the element of chance, randomness or luck. As a consequence, there are neither perfect tests nor treatments. Although not intuitive, the false positive rate of a test—the probability that the positive test is false—varies widely among patients depending on the patient's prior (before the test) probability of having a disease. One of the most accurate tests available is the HIV antibody test. The false-positive rate of the HIV test varies from 3.2% in high-risk for HIV populations to 99.5% in low-risk populations[4]. Thus, if a patient is at low risk for HIV, a positive HIV test is falsely positive in 99.5% of cases. However, when the risk level or prior probability of the disease is not known before testing, the uncertainty of testing vastly affects the decision making for an individual patient. Similarly, among the top ten grossing medications in the U.S., only 1 in 4 to 1 in 25 patients receiving them benefit from the treatment[5]. Thus, some patients benefit, however many do not, and it is uncertain who will respond and who will not. Therefore, in choosing a test or treatment, uncertainty plays a significant role and substantially affects the decision making.

Patients, physicians, and policymakers have been shown to have difficulty with interpretation of the meaning of numbers, especially probabilities. This difficulty is common and has serious consequences for healthcare[6]. A study of 160 obstetricians and gynecologists have shown that 80% of them incorrectly interpreted the risk of breast cancer associated with a positive mammogram. "The majority of them grossly overestimated the probability of cancer," overestimating the risk by almost an order of magnitude 80-90% instead of 10% in the given clinical scenario.[7] US obstetricians and gynecologists also had difficulty in interpretation uncertainty and probabilities related to ovarian cancer screening.[8] The difficulty in interpreting uncertainty and the related probabilities by obstetricians, gynecologists, and other specialty physicians and patients has been shown regarding ovarian cancer screening and other areas.[8, 9]

Because of the uncertainty, the best decision is not the one which results in a good outcome, but one which carefully considers all potential future outcomes, their probabilities and consequences as well as relates to the decision maker's individual preferences[10]. Two approaches emerge to handle uncertainty in medicine. The

traditional approach has been used in medicine from its origins. Yet, recent developments of computational methods and computer capabilities have placed us at a point in time when these advances can revolutionize medical decision making and medicine in general.

The traditional approach to handling uncertainty has been successfully used in medicine for a very long time. For centuries it was based on expertise, experience collected in the process of practicing medicine. Expertise is undoubtedly very effective and saved countless lives. It is especially useful, however, when used with the awareness of its limitations. These limitations of expertise are due to the heuristics we use in intuitive decision making. They were first described by Tversky and Kahneman over 40 years ago, and the latter received for it the Nobel Memorial Prize in Economic Sciences[11]. Heuristics are used by every human being and probably constitute an evolutionary advantage. They allow quick decision making and are usually very effective. However, they may also lead us to make predictable mistakes resulting in a bias in our decision making.

One of these biases is our tendency to overestimate the probability of the events we can more easily retrieve from our memory, the events which are common in our experience and events that are emotionally charged. That is why we all have a biased tendency to overestimate the probability of good outcomes, especially in obstetrics, or specific rare adverse outcomes that have happened to one of our patients[2, 3]. The use of expertise in the handling of uncertainty in medical practice is also hindered by limitations of individual experience. An obstetrician in the U.S. delivers, on average, about 140 patients a year[12]. This means, on average, an obstetrician will experience one cerebral palsy associated with intrapartum hypoxia in 174 years[13, 14], one permanent brachial plexus palsy in 40 years[15], one cerebral palsy due to uterine rupture in 694 years[16, 17], and one maternal death in 38 years[18, 19]. Thus, it is very difficult to collect sufficient experience and use it effectively in estimating uncertainty in medical decision making. The limited possible individual experiences, together with cognitive biases, limits the effectiveness of expertise in handling uncertainty in decision making. These limitations of expertise are well illustrated by words spoken in an area outside of medicine but very well suited to it, by Captain Edward Smith in 1907 “I never saw a wreck and never have been wrecked, nor was I ever in any predicament that threatened to end in disaster.” Five years later, on April 14th, 1912, he was the captain of RMS Titanic.

Mainly in response to the limitations of expertise, as well as the quality and use of evidence, evidence-based medicine was introduced in the 1990s. It was developed with the goal to guide clinical practice by using the results of high-quality evidence, especially in randomized controlled trials (RCTs). This approach was expected to minimize the uncertainty in medical decision making and, in consequence, improve outcomes through consolidation of high-quality-of-evidence in systematic reviews and clinical guidelines. Much progress in medicine has been achieved in the last 30 years thanks to the evidence-based medicine paradigm[20]. However, its limitations and

misapplications have also become evident[21]. Over two decades later, a review of the state of evidence-based medicine shows that many evidence-based guidelines are not based on RCTs and perhaps as few as 11% in some areas of medicine[22]. A review of approximately 3,000 treatments classified 50% of them as having insufficient supporting evidence. Among the other half, judged to be evidence-based, 24% of treatments were considered likely to be beneficial, 7% required trade-offs between benefits and harms, 5% were unlikely to be beneficial, 3% were likely to be ineffective or harmful, and only 11% were clearly beneficial[22].

Perhaps, even more importantly, when the clinical guidelines make recommendations based on RCTs, the recommendations frequently differ between the guidelines even when the same RCTs are cited as evidence for the different recommendations[23, 24]. This is a consequence of the focus of the guidelines' appraisal tools on the internal validity of the referenced RCTs. These appraisal tools focus on the methodology and quality of reporting of the cited studies rather than on their external validity, generalizability, and clinical relevance and applicability besides their internal validity[25]. Therefore, the RCTs may not be applicable to the populations, interventions, and outcomes specified in the recommendations made in the guidelines. Analysis of national clinical guidelines issued by professional organizations from the U.S., Canada, and Europe, showed that of the 338 treatment recommendations made in the nine guidelines, a third had not been based on evidence from RCTs, considered to be the highest quality level of evidence. Another third of recommendations did cite RCTs in support but were found to provide evidence of low quality. The low quality of the evidence in those recommendations was due to lack of applicability of the RCTs to the population targeted by the recommendations, or because the cited trials reported surrogate outcomes rather than the outcomes addressed by the recommendations[25].

RCTs are traditionally considered to be the strongest form of evidence for clinical decision making. However, 20% of all published medical research was shown to have methodological flaws, with RCTs having as many limitations as other studies[26]. Some estimates of the non-reproducibility of the RCTs are even higher. Furthermore, even the highest quality medical evidence is itself uncertain. Analysis of 49 studies, each cited over 1,000 times and published in the leading journals, showed that a third of their findings could not be reproduced by subsequent studies of similar or larger size, or the effect sizes of the subsequent studies were substantially smaller[27].

The RCTs are widely considered the "gold standard" of evidence, and their integration, the meta-analyses, are thought by many as the "platinum standard" of evidence-based medicine[28]. They are considered to be the highest level of evidence, thus the best way to handle uncertainty in medicine. However, despite many tremendous contributions made to the practice of medicine, the RCTs and their meta-analyses have a large number of potentially serious limitations, making them less than optimal to be the sole source of evidence in managing uncertainty in medicine. A. L. Cochrane, the pioneer of the use of RCTs in medicine, warned in 1971, "Between measurements based on

RCTs and benefit . . . in the community there is a gulf which has been much under-estimated"[29].

Today it is evident that treatments used across the medical specialties are consistently less effective in clinical practice than they are reported in RCTs[29-35]. The low effectiveness of medications in clinical practice appears to result from poor external validity or generalizability of clinical trials. Research into the internal validity of RCTs dwarfs the evaluation of their generalizability, and their use in the clinical practice. The RCTs' appraisal tools evaluating their performance, and thus the quality of their evidence usually do not consider generalizability of RCTs, the quality critical to managing uncertainty in clinical practice. In RCTs, only a subset of a population in which an intervention or medication is applied benefits from it[33, 36, 37]. Frequently the benefiting subset is a minority of patients receiving a medication[33]. A still smaller subset of patients may receive a net benefit from medication, the benefit minus harm related to medication[38, 39]. The low generalizability or external validity of clinical trials and, in consequence, low effectiveness of medications in clinical practice is due in large part to heterogeneity of treatment effect (HTE), which has been observed across the spectrum of interventions and domains of medicine[29, 35, 40, 41]. HTE has four sources: heterogeneity of baseline disease risk, heterogeneity of treatment effect, heterogeneity of treatment-related harm and heterogeneity of competing risks, risks related to conditions other than one studied[32, 33, 35, 42, 43].

The limited generalizability of the RCTs is also due to a very narrow selection of study populations in RCTs, and the extensive inclusion and exclusion criteria applied[44]. Thus, the RCTs and the guidelines cannot always be assumed to provide high-quality evidence for the recommendations they make. RCTs and their meta-analyses answer the question "does it work" rather than the question more critical to medical decision making: "in whom might it work." Evidence-based medicine in general and RCTs specifically are substantially limited in handling uncertainty, which is critical in medical decision making and the practice of medicine.

Complex systems

The traditional approaches to handling uncertainty, based on expertise or evidence-based, are shown to have severe limitations. One could argue that these limitations underlie a large part of the health care crisis by enabling abuses of the fee for service system and defensive medicine. The difficulty in managing inherent-in-medicine uncertainty comes from the traditional understanding of the human body and health care as linear cause and effect systems. However, the human body and health care are **complex systems**, systems composed of a vast number of relatively simple components, which intensely interact with each other, and lead to the emergence of unique system behaviors. The complex systems are networks of tightly coupled components, e.g., genes, proteins, cells, etc. interacting in a non-linear manner[45-48]. The unique behaviors of the system emerge from the **interactions** of the simple components. The number of possible interactions grows exponentially with the number of elements in the system. To put this in a proper perspective, a moderate number of

25 elements, for example, risk factors and protective characteristics, could have up to 2^{25} or 33,554,432 interactions among them. If one considers that the human body is made of an order of 10^{13} cells, and each cell contains 42 x 10^{6} protein molecules, the number of potential interactions is genuinely staggering. Interactions, where almost everything affects everything else and does so in a non-linear manner, result in redundancies which make the complex systems robust to failures, such as disease in an individual patient or inability to provide health care to patients in need in a health care system[49]. However, these redundancies also result in multiple ways that failure can occur rather than a single "root cause," leading to failure. Due to that structure, complex systems have many different causative pathways, and thus many different ways an outcome can occur; the phenomenon known as **equifinality**[50]**.** Equifinality, the many different causative pathways and ways an adverse outcome can occur, requires **individualization** of care: *The individualization of decision making, individualization of testing and treatments* (Figure 1). For example, in a predictive model using a combination of 13 continuous and categorical predictors, 80% of the 2 M pregnant women have a unique combination of those predictors - risk factors and protective characteristics.

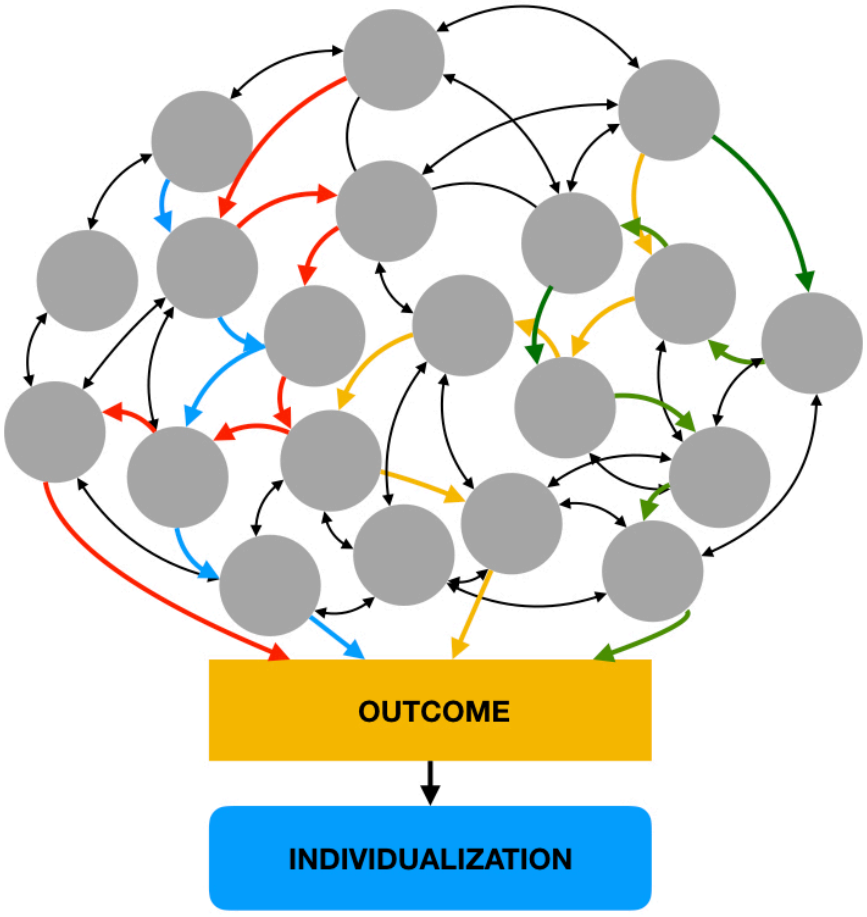

Figure 1. Simplified idealized representation of a complex system, such as the human body or health care system. The system is made of a vast number of components highly interacting with each other. Complex systems have, due to their structure, many different ways or causative pathways through which an outcome can occur; the phenomenon known as equifinality. Equifinality demands individualization as an approach to complex systems such as the individualization of medical decision making.

Thus, any contemplated preventive or therapeutic intervention, test or health policy considered must be based on a balance between individual associated risks and benefits, and an individual decision maker's risk preferences. The human mind is ill-equipped to understand and manage the level of complexity needed for optimal medical decision making. Despite that, these complex decisions are made daily by a broad spectrum of decision-makers ranging from an individual patient to a health policymaker.

Computational models are uniquely suited to handle the intensity of interactions and the uncertainty and complexity associated with them. They excel in modeling complex systems. The word model in this context is a simplified approximate representation of the complex system, which allows analysis of the complex system behavior. The computational part is the mathematical and quantitative representation of the complex system, which prevents ambiguity of the model specification and representation in the form that is executable using a computer. Computational modeling of complex systems can be achieved by using data-driven and/or theory- or knowledge-driven models. In very general terms, the data-driven models require a limited amount of knowledge about the system but a large amount of data. The theory- or knowledge-driven models require a limited amount of data, but substantial knowledge of the system. The computational models predict the behavior of the system, e.g., patients, in the future, and this ability is expressed as external validation of the model. The computational model, which predicts well, also provides a good understanding of the complex system it represents. **Prediction** is an essential component of medical decision making. Because of the inherent uncertainty in medical decision making and complex systems in general, the computational models are stochastic. That means that the models limit the number of possible outcomes by eliminating some possibilities, and giving a higher probability to other outcomes that are among a pool of potential outcomes known before modeling and ones which become a possibility through the use of the model. Familiar examples of computational models are weather forecasts or hurricane predictions.

The computational models **excel in modeling complex systems and at solving these kinds of complex problems**. They are ideally suited to address the critical and complex elements of individualized medical decision making. Computational methods are unrivaled and can play central role in individual predictions and weighing individual, frequently conflicting risks and benefits with their preferences. Computational approaches can account for these types of complexities, aiding optimal medical

decision making for an individual patient as well as simulating a suite of decisions for policymakers.

Computational Medicine can thus be defined as the discipline that uses advanced mathematical approaches to model complex systems along a spectrum from the human body to the health care system. To accurately represent these complex systems, the models need to capture the individuality of health and disease for accurate decision making at all levels. Ranging from the patient to the policy, these require substantial computational capabilities to make accurate decisions. The models can be theory- or knowledge-driven or data-driven, or a combination of these.
To define the future direction of computational medicine, the **Computational Health Conference,** held in Austin, TX in 2018, brought together key stakeholders in health care and experts in computation and health from academia, government, industry, philanthropy, and communities, with the goal of identifying future directions and opportunities for computational health and medicine (Figure 2). In addition to establishing a consensus that advances in computation are creating a new paradigm for medicine and health care, a primary outcome of the conference was a recommendation to create a multidisciplinary Center or Think Tank to research and develop computational solutions to problems of interest for health care stakeholders from academia, government, industry, philanthropy, and communities. Such a collaborative **partnership** would most efficiently use computational methods to exploit data at a deeper level and assure the use of the computational solutions for a range of the stakeholders' interests, from clinical practice to policymaking.

Figure 2. The Computational Health Conference, held in Austin, TX in October of 2018. The conference brought together key stakeholders in health care and experts in computation and health from academia, government, industry, philanthropy, and communities with the goal of identifying future directions and opportunities for computational health and medicine.

Computational models allow to explore a very large number of predictors and many more interactions among them, inherent to every complex system, thus to extract a deeper level of information contained in the data. The deeper, more refined insights allow in consequence better modeling of complex systems by capturing their individual facets and thus allowing individualization of medicine through individualized decision making.

Data are essential, although to a different degree, for computational models to successfully model complex systems. However, having more data on an individual level, higher granularity data, does not necessarily equate to more information. Indeed, there is no perfect data source, neither for traditional nor for computational solutions. Each data source in medicine has its own advantages and limitations. However, computational methods appear to extract more information from the same data than traditional solutions.

DATA

Traditional Sources of Data

Using information to make and inform clinical decisions is innate within medicine. Teaching and personal experience lead to decision making which can be interpreted in light of statistical theory. Bayesian statistical methods estimate posterior risk by combining prior risk with disease associations, which increase or decrease the odds of disease. Clinicians do this every day without formalizing – or possibly even thinking about – the analysis. The experience of sudden onset dyspnea is a possible presenting symptom of pulmonary embolism. The proportional increase in risk (the positive likelihood ratio) will be greater still if it is associated with pleuritic chest pain. The likelihood of the diagnosis of pulmonary embolism will also depend on the prior risk. It will be higher in a pregnant woman with a recent history of long haul travel whereas it will be lower in a febrile but previously healthy child.

Computational methods allow data to be explored at a deeper level by exploring very large number of predictors and many more interactions among them. Hence, complex combinations of positive and negative factors can be combined into a single estimate of risk in a way that an individual clinician would be unable to achieve. Moreover, whereas a clinician may be biased by individual experience of a relatively small number of cases, computational methods attach weights to risk factors which are based on the observed data and can use far more cases than any clinician could ever see. However, there is no perfect data source, either for analytic or for numeric solutions. Each data source in medicine has its own advantages and limitations.

Real World Data

Every data source can be criticized and there is no such thing as a perfect data source: they are all imperfect but they are imperfect in inherently different ways.

Administrative database

The strength of administrative databases is that they tend to have large numbers of subjects and they are available for analysis, i.e. it is usually straightforward to obtain permission for research and the data usually come in a format which is easy to accommodate. This means that rare outcomes can be studied and that the costs of the research are often modest when compared with alternative approaches. There are, however, multiple drawbacks of using administrative datasets. The data are generally of rather low granularity. Hence, outcomes, such as preeclampsia will simply be recorded in a binary fashion with little or no information of the different phenotypes. Key exposure data might also be absent. Consequently, associations may be observed through residual confounding by rather obvious and well recognized confounders. The lack of granularity also affects covariates. For example, a woman may be documented as a smoker. However, it might only be recorded at one stage in pregnancy, the number or type of tobacco might not be documented and there may be no information on whether she subsequently quit smoking during the pregnancy. The other major issue with administrative data sources is the quality. Hence, taking the example above,

not only is the diagnosis of preeclampsia binary, in many cases the definition of preeclampsia may be completely incorrect, being documented as absent in a woman with severe disease and documented as present in a healthy woman. There could also be conflicting information contained in the different parts of patient records.
The strengths of administrative data mean that it is very likely that they will continue to be used for the purposes of research. The appropriate response to weaknesses is their recognition by those generating and using research. The researcher may use multiple sensitivity analyses to determine whether their associations are likely to be true. For example, where there is an analysis of an association with preeclampsia, they might perform exploratory analyses to determine whether the pattern is consistent with a real association. Was the diagnosis of preeclampsia associated with known risk factors, such as nulliparity and obesity? When the outcome was confined to women with other features consistent with the disease, such as preterm birth and fetal growth restriction, were the key observations still present?

Electronic Medical Record (EMR)

The future prospects for researchers planning a career in the analysis of routinely collected data look quite bright and this is due to the increasing use of the EMR. Administrative databases, as described above, have existed since the second half of the 20th century and would have involved records' staff entering data to a dedicated database from a patient's paper case record. However, with the development of the EMR, in some hospitals every piece of information that is held on a patient may be kept in an EMR. This means that the information available is much greater in scope and much more detailed. Moreover, the data are being entered by clinicians who have a much more highly developed clinical knowledge than records' staff. However, research analysis of the data is a secondary purpose of the EMR. As the primary purpose is to facilitate the delivery of and billing for care, the data are entered prioritizing this end. Moreover, although the clinical staff will have greater knowledge, data entry is the means to an end and not their primary purpose. Hence, when compared with records' staff, they may apply definitions less consistently and incompletely. A great strength of analysis of the EMR is the access to observed numerical data. This includes vital signs (e.g. temperature, blood pressure, and respiratory rate) and the result of lab tests, (e.g. biochemistry, hematology, and microbiology). Hence, if studying preeclampsia, a researcher might hesitate to accept the presence or absence of the diagnosis being documented in the EMR as defining the presence or absence of the condition. However, considering the ACOG Guidelines[51] for definition, every single blood pressure measurement recorded during a stay could be analyzed to define hypertension, every point of care and laboratory analysis of urine could be used to define proteinuria, and every laboratory test performed during the admission (such as creatinine, platelet count or alanine transaminase) could be used to define the renal, hematological or hepatic features (respectively) of severe disease.

The practical utility of EMR data also depends on other issues. Whereas administrative databases are typically available in a simple spreadsheet format of columns and rows, although the very large databases may have a more complex structure, EMR data will

reflect the complexity of the clinical environment. Some members of a cohort of pregnant patients may have 10 measurements of blood pressure, others may have 200. The researcher needs to consider whether they wish a raw data dump, or whether the information is pre-processed. Whatever the case, additional resources will be required to format the dataset prior to analysis, when compared with an administrative dataset. Statistical power is likely to be better than in many research studies, as the entire population attending a hospital will, over a short period of time, generate the sorts of numbers of patients that would be extremely expensive to study in a prospective fashion. However, outcomes such as perinatal or maternal death may still be too infrequent for the analysis of a single institution to be informative. Perhaps one of the most promising avenues is the analysis of a common EMR used by a network of hospitals, perhaps covered by the same provider, or where different providers have used the same EMR. For example, a national randomized controlled trial in the UK is going to collect outcome data from a widely used neonatal intensive care EMR to ascertain outcomes[52].

Research Studies

Study design is key when assessing the evidence around a given belief in medicine. This has led to the widely adopted use of "levels of evidence" (Figure 3). However, the prioritization of study design above many other important issues is open to criticism[53]. For example, for some outcomes, the main challenge might be to perform a study that has a large sample size, as the outcome in question is rare, which is true for the most important outcomes, such as moratlities or severe morbidities. Randomized controlled trials are biased to report negative results for rare outcomes as conducting a trial which is powered is prohibitively expensive. Similar arguments apply for remote outcomes, due to the expense of long term follow up. Hence, the "pyramid of evidence" could be an unhelpful metaphor, as it oversimplifies a complex question.

Observational studies

There are multiple observational study designs and a description of these is outside the scope of this review. The key issue around observational designs is that the exposure of interest is observed, but not determined by the investigator. This is in contrast to the RCT (described below), where the exposure is applied experimentally. The major weakness of observational study designs is that associations between the given exposure and the outcome are not causal but are related to their mutual dependence on a third characteristic, a confounder. Controlling for confounding can be attempted statistically, for example, using multivariable statistical models. However, the concern is that an association which is observed after statistical adjustment for the measured potential confounders could be explained by an unmeasured confounder. For this reason, the associations described in observational studies are interpreted cautiously. The reality is, however, that many major questions can only be assessed using observational studies. Sometimes, this is because women will not agree to be randomized to a trial. Virtually the entirety of the evidence around vaginal birth after cesarean (VABC) is based on observational data. A group who attempted a large scale trial RCT of VBAC versus planned repeat cesarean delivery in 14 Australian maternity

hospitals only managed to recruit 22 women out of a cohort of 2,345 (i.e. <1% of those recruited)[54]. Other questions will only ever be addressable using observational studies because an adequately powered trial is prohibitively expensive. For example, there is strong evidence from observational studies that there is an increased risk of perinatal death of the second twin at term, but not preterm[55]. Observational studies also indicate a reduced risk of perinatal death of twins with planned cesarean at term[55]. However, a multicenter, international RCT was only able to study about 1,500 term twin births[56], less than a quarter of the required sample size to study the effect of planned cesarean on the risk of perinatal death[57]. Interestingly, a recent re-analysis of the Twin Birth Study has shown that planned cesarean delivery reduced the risk of a composite adverse outcome, including death, when delivery occurred at term but not preterm (as predicted by observational studies)[58].

Randomized controlled trials

There is no doubt that an adequately powered randomized controlled trial, conducted in a methodologically rigorous fashion (e.g. prospective registration, pre-specified primary outcome, pre-defined analysis plan, overseen by independent steering and data monitoring committees) provides the strongest evidence in relation to the effects of interventions in medicine. Furthermore, meta-analysis of multiple such trials, conducted in different settings, will generate more precise estimates of effect size and assess whether the intervention works consistently in different settings. However, a minority of clinical decisions in obstetrics and gynecology that are made using such an evidence base.

An increasingly recognized concern is that we may end up making clinical decisions based on meta-analysis of randomized controlled trials but the decisions made may be flawed due to issues with the evidence base, despite being composed of RCTs. As recently stated by the Editor of The Lancet “But what if the astonishing energy, commitment, and productivity of the systematic review community are poisoning rather than nourishing medical practice?”[59]. The issue being addressed was concerns about the quality and accuracy of many of the small trials included in systematic reviews. However, even well conducted trials have the potential to mislead. The women recruited to a trial might not be representative of the general population: this undermines the external validity of the conclusions. RCTs are difficult and expensive, hence, the number of women recruited may have been limited by cost meaning that only relatively common outcomes could be studied. Hence a treatment might be recommended or not based on its effect on mild adverse outcomes which are common when it has the opposite effect on severe adverse outcomes which are rare. Cost may also limit the duration of follow up. Hence, an intervention might be recommended on the basis of a short term benefit but in the absence of evidence about its long term effect.

Further issues relate to type 1 and type 2 statistical error. Whereas best practice recognizes that reliable conclusions from RCTs can only be drawn from the pre-specified primary outcome, “p-hacking” still occurs. Multiple hypothesis tests make it more likely that a null hypothesis is incorrectly rejected, type 1 error[60]. Conversely, type 2 error is the major concern in relation to statistical power. If the sample size is too

small, a study will be biased to produce a negative result. In an ideal world, clinicians would look at the 95% confidence interval of the effect and understand that no safe conclusion can be drawn. However, with an over-emphasis on p values, the statistically uninformed may equate absence of evidence with evidence of absence and utter "evidence-based medicine's six most dangerous words"[61]: "there is no evidence to suggest that..."

In summary, clinical decisions involve drawing conclusions from a large body of evidence. For most clinical decisions, the task of relating the research evidence to making a given decision is highly complex. Analysis of routinely collected data, whether administrative or EMR, has the advantage that the entire population can be studied, but has the drawback that the approach is much more complicated and the output can be difficult to assess. RCTs are attractive by their relative simplicity and their experimental design. However, considering complex questions about recommending a given approach in a simplistic manner, e.g. study design is the only major concern, has the capacity to lead to widespread harm.

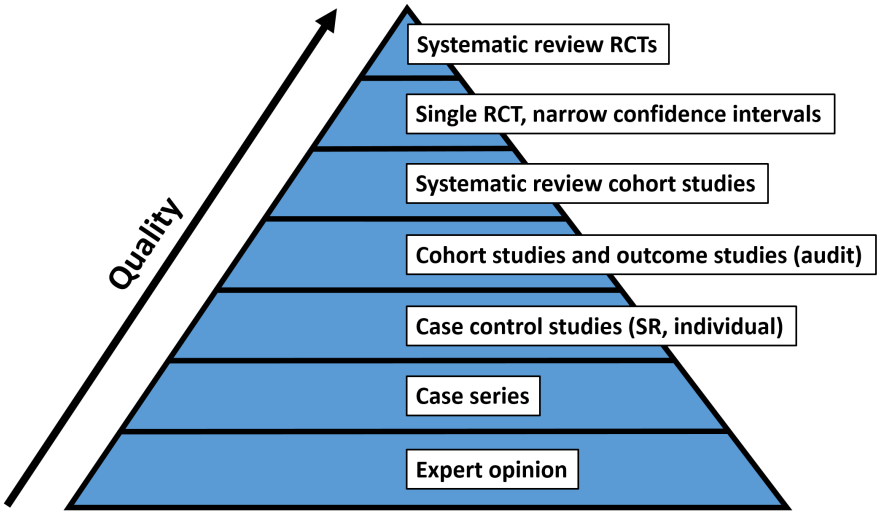

Figure 3. Schematic representation of the "levels of evidence". RCT denotes randomized controlled trial and SR denotes systematic review.

Novel Sources of Data

Digital phenotyping

Digital phenotyping is an all-encompassing term for the trail of digital data that people leave behind in their daily lives – interactions with the internet, social media, and technology. This data is largely untapped and has significant potential for use in the health care industry. People generate an enormous amount of digital data each day, and this moment-by-moment computation of an individual's phenotype measured in-

situ from ubiquitous personal devices has enormous potential to revolutionize the way we understand and make sense of health and health-related conditions[62].

81% of Americans now own smartphones, and this rate has dramatically increased over time[63] (Figure 4). The ubiquity of smartphones makes them ideal to collect detailed patterns of true behavior from individuals in an objective manner. Additionally, passive methods for collecting data – those mechanisms that automatically collect, with the individual's consent, but without the need for explicit input – provide a continuous means of collecting data in the background and allow for a more fine-grained collection of behavioral, health-related, environmental, and lifestyle data[64]. This data, in combination with traditional clinical data, provides a powerful tool to understand and develop a digital phenotype of both health and disease, at the individual level but also across a given population (Figure 5). With rapid advances in the broad availability of machine learning and the application of numerical simulations, digital phenotyping holds tremendous promise to provide a plethora of insight not previously possible. Novel approaches such as digital phenotyping allow wider applications of N-of-1 trials or studies and are more effectively performed using computational methods. By connecting aggregate multi-level and multi-scale clinical, biomedical, personal, social, contextual, environmental, and organizational data and using it to individualize decision-making in health and medicine, we would get not only a more informed picture of an individual's health, but also a greater promise for transforming and optimizing health care for populations as well.

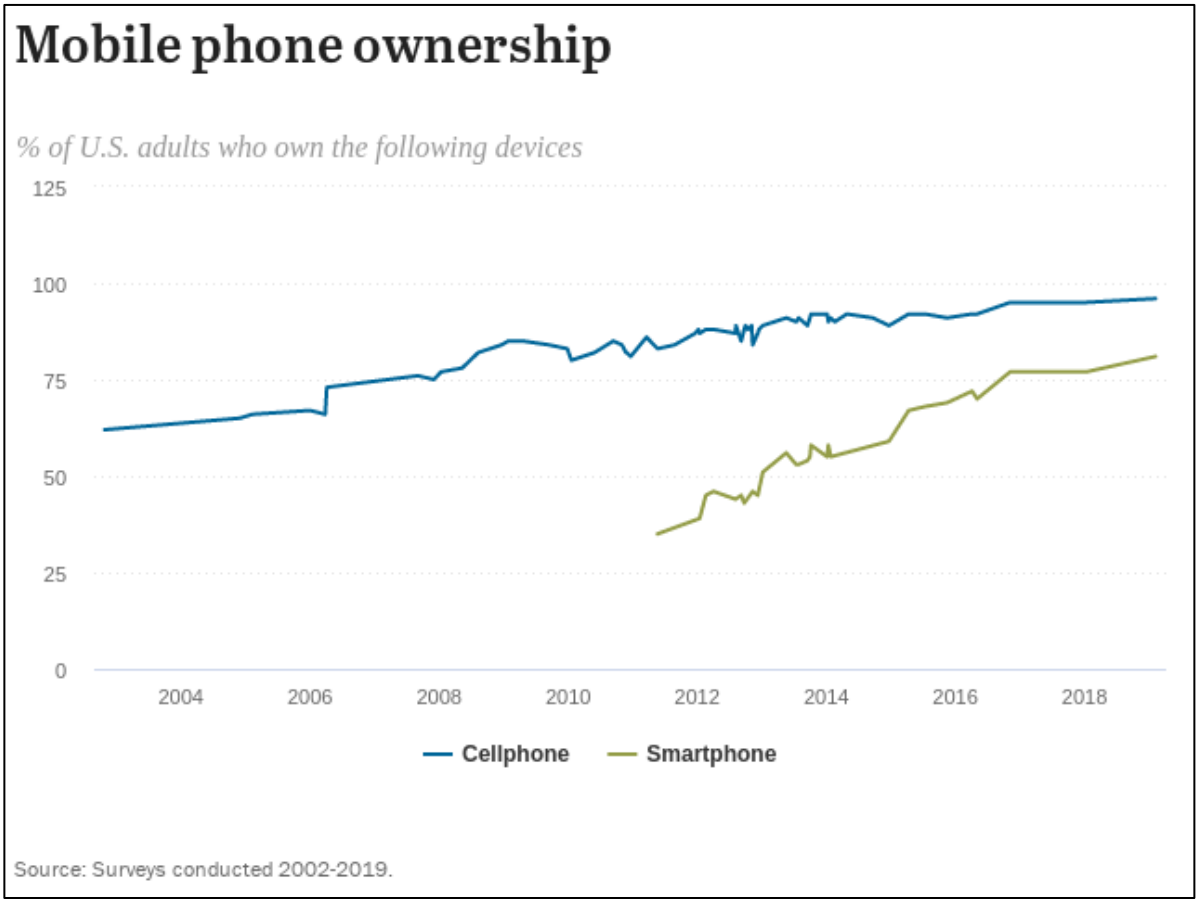

Figure 4. Proportion of U.S. adults who own cellphones and smartphones.

Figure 5. Monitoring interface of digital phenotype data streamed real-time from a patient's cell phone to the supercomputer.

Wearables and internet connected sensors

Current and emerging wearable and internet-connected devices/sensors are poised to address many deficiencies of traditional data sources by enabling the longitudinal collection of contextually rich, multifaceted, and individual-level data at unprecedented scales. The use of wearables or smart devices (e.g. activity trackers, smartwatches, glucometers) for health and wellness is rapidly proliferating among consumers from an estimated 9% in 2014 to 33% in 2018[65]. They also continue to be incorporated into clinical practice as a means of remote patient monitoring (e.g. chronic disease management, acute health event detection) and population based health research(4). A detailed review of these devices is beyond the scope of this article and we refer the reader to Dunn et al.[66] for a thorough discussion of health-related applications. The diverse data generated by sensors embedded in wearables or smart devices can report on, for example, an individual's mobility and physical behaviors using accelerometers, gyroscopes, and GPS; physiological parameters like heart rate, temperature, and oxygen saturation; biological analyte concentrations like blood glucose measured from continuous, internet-connected glucose monitors; and real-time air pollution exposure via personal air quality monitors[67]. Furthermore, an individual's social behaviors and mental health may be assessed, for example, through the analysis of phone call and text message logs (i.e. frequency, duration, incoming/outgoing)[68] or through psycholinguistic analyses of user-generated social media content[69]. These technologies have been used in a variety of health domains including weight management, metabolic and cardiovascular disease, maternal and neonatal care, sleep quality and assessment, and neurology, among others[66, 70]. While it is likely that the next decade will see networks of wearable devices being routinely used in medical practice, there are a number of challenges to be addressed.

These include, but are not limited to, the development of a robust cyber-infrastructure capable of the real-time fusion and analysis of disparate data streams and flexible enough to accommodate the rapid pace of consumer technology development; development of analytical methods and algorithms to extract or discover signatures (i.e. digital biomarkers) or leading indicators of health or disease status from 'real-world', "messy" data; execution of large randomized controlled trials to assess efficacy of using these novel data sources in treatments or interventions; and interoperability with existing electronic medical records and medical decision-making workflows.

Contrary to popular belief, the data or "big data" itself is not the focus of computational medicine. Its focus is the modeling of complex systems such as the human body or the healthcare system. Data in this context is an essential prerequisite to building a model, but as highlighted in the section on the theory-based model, a large amount of data and especially "big data" is not a necessary condition for those models[71].
Second, currently, data gathering is primarily driven by convenience and is empirical in nature. Such an approach must necessarily lead to many areas where data is either not collected or not useable for model development. It has been proposed that the models, especially the computational models, should define what data is needed and to direct the design of experiments allowing better model development. The area where such an approach would be especially useful is genomics[71, 72].

MODELS

However, even when the data contains the necessary information, it is by itself insufficient to individualize decision making and medicine in general. Indeed, an accurate model of the complex system has to be developed, a model which captures the equifinality, the multiple causative pathways in the system leading to the same outcome, requiring individualization of medicine. The capture of equifinality and hence, enabling of individualization of medicine can be achieved by data being used by computational models, theory- or knowledge-based models, data-driven models, or models that combine both approaches. As mentioned earlier and using broad strokes, the data-driven models require a limited amount of knowledge about the system but a large amount of data. The theory- or knowledge-based models require a limited amount of data, but substantial knowledge of the system.

Theory- or knowledge-based models

What is a "model": a mouse model for laboratory experiments that try to mimic human physiology, a toy airplane model, a model community, …? No, the notion of a model addressed here is a mathematical construct that represents an abstraction of physical phenomena described by a scientific theory – a mathematical characterization of a set of inductive hypotheses, often based on observation and moving towards generalized conclusions, put forth to explain events that occur in the physical universe. Such scientific models thus form the essence of the second pillar of science, along-side

empirical observations, as a fundamental source of knowledge. To qualify as a scientific theory, or as a meaningful model based on theory, the theory must be "falsifiable", according to philosopher Karl Popper – that is, it must be capable of being contradicted and abandoned if predictions contrary to theoretical predictions are observed[73]. Today, we deal with computational models to make predictions of the behavior of complex systems – they are corruptions (discretizations) of mathematical models constructed so as to be implemented on digital computers. The selection of a model to describe a class of physical realities is the most important and difficult component of predictive science. Models involve parameters that are usually unknown and can be random variables; the calibration of model parameters, the adjustment of the model parameters for the model to better reflect the modelled complex system, requires the acquisition of data which can be noisy and expensive to access, and the discretization of the model to produce a viable computational model introduces additional uncertainties. The quantification of these uncertainties is essential for reliable model predictions.

Science and scientific prediction without models are meaningless. It must be emphasized that the development of predictive computational models is critical to the advancement of science; it is a fundamental challenge that must be met in all areas of science, medicine, and technology.

The construction of predictive computational models in medical science and practice could lead to one of the most important developments in human history. For example, the development of predictive computational models of cancer[74] and of the effects of various cancer therapies for specific individuals, where patient -specific observational data is used to calibrate and validate models, would make possible breakthrough effective and non-invasive treatments of the disease, revolutionize medicine worldwide, and forever enrich and expand the scope of medical science.

Data driven models

Data driven models are those for which predictions are performed by examining relationships between a number of available state variables (predictors) related to a particular quantity (outcome) of interest. Unlike theory-or knowledge-based models, explicit knowledge of the underlying physical, biological, and psychological mechanisms impacting the outcome is not a pre-requisite. Instead, the goal is to develop predictive models that infer relationships directly from available data. Fundamental to the development of these models is the availability of sufficient data containing input features (predictors, confounders, and mediators) which have observed influence (either directly or indirectly) on the outcome of interest. Historically, data driven models used in health care have been fairly simple via the deployment of regression models leveraging a relatively small number of input parameters. These regression models typically have very modest computational requirements. More recently though, more advanced machine-learning models are being exploited in a variety of health-care areas. Relevant examples of modeling algorithms in this space

include feature detection/pattern matching (useful in imaging/radiology), clustering (the detection of like groups and structures), and classification (formulating predictions into a predefined set of relevant outcome classes). Clustering is an example of unsupervised learning where labeled data (data with defined outcomes) is not required as an input, whereas classification is an example of supervised learning that requires a set of data with known outcomes that serves as the basis for training the classification algorithm. Performance and range of applicability of these machine-learning models often improves with larger amount of data, and consequently, these types of data-driven models can lead to more demanding computational requirements requiring parallel processing and high-speed I/O subsystems during their formulation, particularly for model training in supervised learning approaches. While very promising in a variety of industries, there are challenges in the U.S. health care industry related to development and adoption of machine-learning models. Examples of these challenges include data inconsistencies and limited widespread availability of high-quality data for model development due to the proprietary nature of EHR systems, lack of understanding and exposure to more advanced data analytics within the medical community, and difficulty in disseminating machine learning results by health care professionals which can be perceived as "black-box" models.

Combined theory- knowledge-based and data driven models

The proponents of the two types of computational models, the theory- or knowledge-based models and data driven models, often argue about their relative merits, but experience indicates both are very important in computational medicine. An example is illustrative.

HeartFlow, Inc., is a pioneering computational medicine company providing individualized, noninvasive diagnosis of coronary artery disease to determine whether or not coronary artery narrowings are obstructing blood flow and would benefit from coronary artery stenting or coronary bypass surgery. The primary source of patient data is a coronary CT scan, which is uploaded by way of a secure web-based interface. The first aspect of modeling is construction of an individual quantitative three-dimensional anatomic model of the aortic root and coronary arteries. This step,- performed using data-driven modeling, specifically deep learning, is referred to as "segmentation" and is particularly challenging. Construction of a high-accuracy geometric model of the coronary vessels on a patient-by-patient basis, even with the assistance of computer algorithms, may take an inordinate amount of time and is prone to error. However, due to compilation of an enormous data trove of coronary trees of over 50,000 patients Deep Learning algorithms could be trained to develop the data driven models in order to more quickly and accurately segment coronary arteries of new patient data. Data driven models, developed using machine learning, e.g. deep learning, are very effective for visual tasks and constructing geometry from a CT scan is in essence a visual task.

Once the coronary arteries are segmented, an analysis of blood flow is performed through theory- or knowledge-based modeling. The Navier-Stokes equations of fluid

dynamics theory are employed because they are capable of accurately representing blood flow phenomena in the coronary arteries. Additional data is not needed to establish this fact. The Navier-Stokes equations are based on the foundation of Newtonian Mechanics and have been corroborated through untold numbers of physical experiments over hundreds of years. Very accurate, efficient and robust computer algorithms are employed for their solution. The end result is a precise prediction of flow velocity and pressure in the coronary arteries. Clinicians are particularly interested in Fractional Flow Reserve (FFR), which is the pressure drop across obstructed regions in the coronaries caused by disease under conditions of maximal coronary blood flow (hyperemia). This predicts whether or not coronary revascularization is needed and is the gold standard for guiding treatment. During cardiac catheterization, hyperemia is induced by administering a drug (adenosine). HeartFlow analysis simulates this hemodynamic condition computationally, thus making it possible to non-invasively determine which patients should be treated medically and which should be sent for coronary angiography and possible revascularization. The computational model can also be manipulated to simulate stent implantation with prediction of the potential outcome of stenting before the procedure is performed. The benefits of HeartFlow's synthesis of theory- or knowledge- based and data driven modeling have been demonstrated by the Platform Trials[75]. Among patients with planned invasive coronary angiography (ICA), 73 percent showed no significant blockage or obstruction and in 61 percent of patients, the use of computational modeling resulted in the cancellation of a planned ICA. After one year, none of the 117 patients who had ICA cancelled had suffered an adverse clinical event. There are over 1 million ICAs in the U.S. each year, and over 2 million in Japan, and a similar number in Europe. The cost of an ICA in the U.S. is over $12,000; about 2/3 seem to be unnecessary and can be eliminated, representing potential savings in billions of dollars. In summary, there are well established and highly-reliable theory- or knowledge- based models that can be effectively used in computational medicine, but where these models do not exist, we can utilize powerful data driven modeling approaches.

Computational models, as any other models in medicine, are developed to be clinically useful to make optimal medical decisions and thus improve patients' health. To improve decision making, the models need to enable both individual decision making and determination of individual net benefit for contemplated tests, or treatments[76]. To aid in clinical decision making and especially in individualized decision making, the models need to have high discriminatory power and be well calibrated[77]. Discrimination is the ability of the model to predict a higher probability of the event in patients who will ultimately experience the event – outcome than in patients who will not have it[78]. If a model always predicts a higher probability of the event in patients who have it than in patients who do not experience it, then the discrimination of this model is perfect, and its measures, Area Under the Receiver Operating Characteristic Curve and c-statistics are equal to 1.0. However, the relationship between true positives and false positives for different cut-off points of the model predictions those statistics describe has to be balanced. Thus, choosing cut-off, which maximizes true positives, leads to an increase

in false positives and vice versa[78]. Calibration, on the other hand, is the ability of the model to predict probabilities of the event (the absolute risk) that are in agreement with observed frequencies of the event[78]. In other words, if a well-calibrated model predicts risk of the outcome of 60% in 10 women, six of the ten or 60% of them will experience the predicted event.

Good discrimination is an essential first step without which the model is generally not useful for decision making. However, good discrimination is insufficient for effective decision making. The measures of discrimination assume an equal value of sensitivity and specificity, thus of false negatives and false positives. In real life, however, the consequences of false negatives (e.g., missed diagnosis) are more severe than consequences of false positives (e.g., unnecessary tests)[76]. As a consequence, a model with good discrimination can still have an unacceptable rate of false negatives – missed diagnoses[76]. A poorly calibrated model can lead to a situation where an individual with a high risk of disease has assigned a low probability of it occurring and thus misses the opportunity of effective preventive intervention[76]. Although both the discrimination and calibration are critical for the clinical usefulness of the model for medical decision making, they are rarely both provided in the literature[78]. A systematic review in cardiovascular medicine showed that discrimination was reported in 63% and calibration in 36% of the studies[79]. Good discrimination and good calibration of the computational model or any other model is necessary for the model's clinical usefulness for individual patient and individualization of medical decision making[77, 80]. The rarer the event the model predicts, the more difficult it is to achieve good calibration. This is because the model is much more likely to be right by predicting the overwhelmingly more common non-event, and thus the prediction of higher risk of the event is underestimated. However, the outcomes of highest interest are the rare most severe morbidities and mortalities, rather than more common proxy outcomes. Finally, because decision making is about the prediction of future outcomes in individuals and populations, the model's discrimination, calibration, and usefulness in decision making can only be accurately evaluated in the process of external validation[81]. That is in the population of individuals whose data were not used in model development and who are from a population that would be a potential target population for the clinical use of the model. Therefore, external validation should be the final adjudicator of the model's performance and clinical usefulness for decision making.

Clinical practice can only be informed by very good models that well represent the complex system of the human body in health and disease accurately. The data are critically important but are, conceptually, insufficient to the development of such models. Building meaningful models requires a much broader analytic and quantitative medical expertise, much beyond empirical inputs. Novel sources of data, the signals of wearables, digital phenotype data including environmental or social media data, etc. carry a promise of improving the accuracy of the models and, as a consequence, individualization of health care that utilizes those models. However, at this point, it is not clear which sources of data and to what extent will inform the development of the

models in general and computational models specifically. In short, more data does not mean more information and, thus, better health care.

Computational models offer an opportunity to integrate the different sources of multimodal data into actionable information, which will inform the clinical practice. Computational models were shown to provide real insights, for example, into breast cancer therapy[82]. However, although promising, the clinical usefulness of those models has not yet been demonstrated, including those proposed in obstetrics and gynecology[76, 80, 83, 84].

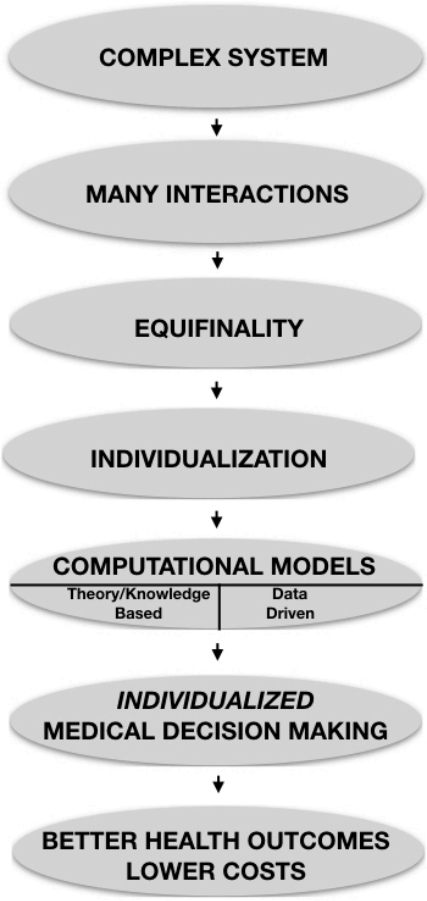

Figure 6. The role of computation in medicine. The human body and the health care system are complex systems, networks of highly coupled components intensely interacting with each other. These interactions give those systems redundancy, thus robustness to failure and, at the same time, equifinality, many different causative pathways leading to the same outcome. The equifinality demands individualization of medical care, which is urgently needed. Computational models excel in accounting for a very large number of interactions, thus in the modeling of complex systems, and hence enable in the individualization of medicine. They have the potential to enable

individualization of medical decision making and, in consequence, better health outcomes and lower costs.

MEDICAL DECISION MAKING

As we have argued in the beginning, the human body, as well as the health care system, are complex systems. Those systems are made of a very large number of intensely interacting elements forming highly coupled networks with a vast number of feedback loops and redundancies. This structure makes complex systems robust to failure, but also due to equifinality, difficult to predict their behavior. Due to this pervasive and inescapable uncertainty, medicine is, in its essence, decision making under uncertainty. The decisions about tests to be performed and treatments to be administered range the entire spectrum of health care from decisions regarding individual patients to policymaking. Traditionally, decisions are being made based on experience and traditional evidence from research studies, both limited by a number of factors including limited possible personal experience of a single physician and average effects reported by the clinical research studies, respectively. In complex systems, characterized by many causative pathways, basing the decision on an average effect in a population of patients is limiting the optimality of the decision. Since medications are only effective in a relatively small fraction of patients receiving them[5] and tests' predictive performance depends strongly on the patients' characteristics defining their pretest risk of the disease tested for[4], decision making has to be individualized to be optimal. Specifically, it has to be optimal under the condition of uncertainty, optimal when the decision is being made prospectively and not after the outcome is already known,[10] and for an individual or a population of individuals, and not for a population average. However, to individualize medical decision making requires accounting for very large numbers of constellations of risk factors and protective characteristics. The number of those unique combinations grows exponentially. For example, just 5 characteristics with 3 categories each would result in 243 unique combinations of those characteristics potentially affecting the individual outcome (Figure 6).

Traditional methods of analysis are limited in handling this type of complex data. However, computational models excel in modeling the behavior of complex systems, enabling individualization of decision making. The individualized medical decision making is dependent on the individual prediction of outcomes, individual weighing of probabilities of outcomes and individual (patient to policymaker) preferences, and individual risk communication (Figure 7).

Individualized prediction of outcomes

"Prediction is an essential feature of non-arbitrary decision making"[50] and individual prediction is an essential feature of the individual decision making, critical in medicine. While the perfect prediction of behavior of complex systems may never be achievable, computational models can eliminate, from a pool of potential alternatives, ones which

are inconsistent with the data for an individual and assign individual probabilities to the remaining alternatives[50]. This probabilistic approach is very effective in real-world individualized decision making, especially when we accept that deterministic solutions (0% or 100% probabilities) are not possible in predicting the behavior of a complex system such as human body[85]. Perfectly accurate tests or perfectly effective interventions are not and are not likely to ever be available.

Decision theory: weighing individual probabilities of outcomes and individual (from the patient's to policy maker's) preferences

Building models of individual optimal decision making under uncertainty in health and medical care is of tantamount importance. The modeling cornerstone, besides the individual outcome prediction, is to describe the preferences towards uncertainty, capturing individuals' attitudes towards risk, loss, incomplete information, and other ingredients that affect personal choices, both statically and dynamically[85]. Building risk-preference models is a challenging task as similar notions in economics (expected utility theory, behavioral finance, bounded rationality, rational inattention, etc.), albeit foundational, cannot be directly applied due to how detrimental certain risks might be for the health and wellbeing of the patient. Indeed, traditional risk criteria, based on averaging and smoothing formulations, are not very suitable to model loss aversion, fear, prudence, impatience and other (frequently, acute and also path-dependent) sentiments arising in the course of a medical treatment. Furthermore, traditional criteria are typically one-dimensional and, thus, cannot capture the multi-attribute risks a patient faces.

An additional challenge is how to solve the associated stochastic optimization problems. Indeed, such complex risk preferences criteria often give rise to "time-inconsistency", a well-documented phenomenon even for financial and insurance risks. These problems are very hard to solve because all classical optimization approaches fail and new techniques, both analytical and computational, need to be developed. In addition, these problems need to be analyzed "in real time", for model decay always occurs and adaptive optimization criteria need to be incorporated to capture incoming information.

Overall, individualized medical decision making under uncertainty in dynamic settings is a wide-open area with a plethora of new research directions. While, as mentioned above, modeling patients' risk preferences will borrow considerably from fundamental notions in financial economics, building *both* sophisticated dynamic models and solving the associated stochastic control problems present many challenges but, at the same, a very fertile ground for both cross- and interdisciplinary collaborations. Furthermore, there is a pressing need to support such developments as, at present, there is a rather sizable discrepancy between the sophistication that exists in medical science models and the simplicity, if not absence, of evaluation criteria from the patients' point of view.

The literature and evidence in the area of decision theory in medicine is very limited and new computational and computer science based approaches promise to offer progress in this area.

Individualized risk communication

However, individual prediction and weighing of individual probabilities of outcomes and individual preferences are insufficient for individual decision making in medicine. Those individual probabilities and preferences have to be appraised and communicated efficiently to make the effective decision making. The dominant way risk communication has been viewed is as an individual-level perception. Uncertainty is presumed to drive risk, and it's perception exists when an individual perceives information to be unavailable, inaccessible, or inconsistent[86]. Most theories of risk communication conceptualize risk as an individualized perception (e.g., Planned Risk Information Seeking Model: PRISM)[87]; Theory of Planned Behavior (TPB)[88],Theory of Motivated Information Management (TMIM)[89], Extended Parallel Processing Model (EPPM)[90], Health Information Acquisition Model (HIAM)[91], Risk Information Seeking Model (RISP)[92], and the Comprehensive Model of Information Seeking (CMIS)[93]. In addition to the view that risk communication is an individual-level construct, it is important to realize that perceptions of risk are influenced by others—or socially constructed[94, 95]. Past research has acknowledged this in several communication models (e.g., subjective norms that influence individual behavior), but these perceptions are still measured on an individual level. A more recent argument is for health-related risk studies to think of risk communication as the "exchange of information among individuals, groups, and institutions related to the assessment, characterization, and management of risk"[96]. This is an approach that resonates with health care providers because they are naturally part of the risk communication process. However, the important others are expanding, and this is especially relevant as mobile and social media become a part of the fabric of our society. Specifically, individuals can make their risks visible to others, by posting on social media or sending a photo/text to a trusted friend of family member, and others can directly respond, thus influencing how individuals internalize and potentially act regarding their own risk. Thus, risk decision making is a combination of individualized perceptions and influences of others around them.

Medical decision making relies on the understanding of potential outcomes and their probabilities. A rational decision would theoretically try to maximize the probability of a positive outcome while minimizing the risk, the probability of an adverse outcome. Of course the reality is more complex. There are many factors that impact medical decision making that may cause a patient to choose a more risky procedure in order to achieve a desired outcome. The preferences, values and biases, of the individual patient have a large influence on the perception of risk, the value of potential outcomes, and ultimately the individual decision making. For example, a patient may choose to deliver the baby at home rather than in the hospital even though she knows that the risks for home delivery are higher[97]. Here it is the personal preference and the value placed on home delivery that cause the patient to place less importance on the

associated risk. Knowing these individual biases can help the clinician formulate an effective communication strategy. For instance, knowing that the patient prefers home delivery, it would be prudent to ensure the patient truly understands the associated risks of home delivery in particular if she has other risk factors. Computer game technology has been used in the past to educate patients about the risks and possible outcomes of various screenings and procedures. In the future, immersive game technology will be used to automatically determine how a patient places value on certain outcomes as well as learn what bias the patient has regarding associated risks. The use of virtual environments will allow the patient to be less inhibited about their responses to questions asked through interactive scenarios. The game itself will use these responses to adapt the presentation of information regarding the risks and outcomes of the procedure in question. Ultimately the decision lies with the patient, but with improved communication provided by computational solutions, the clinician and patient will be able to formulate a plan that meets the bias and value structure of the individual patient and minimizes the risks involved.

The individualized decision making, made possible by the computational modeling of complex systems, has the potential to revolutionize the entire spectrum of medicine from individual patient care to health care policymaking. In patient care, computational models can enable individual decision making based on the patient individual net benefit of contemplated tests and interventions. In making decisions on a strategic level of a hospital, a system of hospitals, or on the level of government policy, the computational models enable individualized decision making for a population. Those individualized decisions on the population level are based on the benefits of individuals comprising the population rather than based on the average benefit of the population. This approach allows applying tests and treatments exclusively to individuals who receive net benefit, in whom benefits outweigh the risk, rather than to all individuals in the population regardless if they do or do not receive the net benefit. Computational models enable individual medical decision making that can transform the medicine and the health care system, both in urgent need of a substantial disruptive transformation.

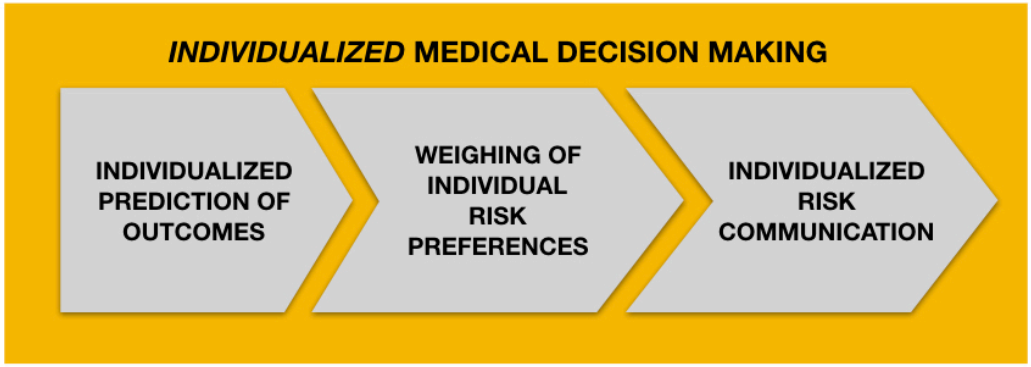

Figure 7. Individualized decision making in medicine can be thought of as composed of three consecutive parts: prediction of the individual probability of outcome; weighing of those individual probabilities of outcomes and individual patient preferences for outcomes; communication of the risk, the probability of the outcome and the burden of this outcome.

FUTURE OF COMPUTATIONAL MEDICINE: HEALTH CARE STAKEHOLDERS' PERSPECTIVES

The substantial potential benefits of computational medicine can only be realized if they solve problems important to health care stakeholders. This will ensure that the computational medicine solutions will be implemented and thus could transform health care. The potential to transform health care exists at all levels of the system, at the hospital, hospital system levels, to government policy. This is because health care is a complex system, which can be thought of as a highly coupled network system in which constituents affect each other[46, 48].

Academia

From an academic standpoint, it is an extremely exciting time to be investing in research and educational programs in Computational Medicine. The confluence of computing power, scalable algorithms, availability of patient data, and increased understanding of physical mechanisms, positions us for a revolution in the way medical advances are achieved, as well as in the way medicine is delivered. The methodological foundations for concepts such as a "digital twin" are being laid in fields as diverse as aviation and medicine, built upon synergistic combinations of predictive physics-based models and data, and providing an enabling technology for achieving asset/patient-specific data-driven decisions. The research and educational programs that will enable this revolution must be highly interdisciplinary in nature. The fields of artificial intelligence and data science will play a big role, particularly through the use of machine learning methods, but we cannot lose sight of the critical importance of physics-based and mechanism-based modeling. Educational programs in Computational Medicine must blend these perspectives, training students at the interfaces of mathematical modeling, computing, data science, and medicine. Partnerships have never been more important, including the sharing of data and digital infrastructure -- both partnerships among units within the university system and partnerships across universities, government, and health care providers.

Hospital system

In health care we are awash in data and information. But, it is often in disparate systems, discrete data, and largely doesn't tell the enire "story" of the patient. With computer technology and the introduction of the Electronic Medical Record we have lost sight of the entirety of patient care, invested more in small, discrete data elements to enter into the system, often to never be harvested. Over the decades we have stored more patient-specific data than ever, but at what cost? Very few, if any health systems

maximize the use of data to truly improve quality, patient safety, health outcomes or to ultimately improve the health of the communities we serve. We have the ability to make profound changes in how care is delivered. Given the current state of health care, the variability in care delivered and continued rise in overall cost of care, we need a dramatically new approach, and the data we have been feverishly storing, may hold the answer to the changes we need.

The discrete data inputs, when provided to computational models, may lead to information which results in better decision making, more decisive and coordinated care, and ultimately, better outcomes for patients and communities. With computational models available, and more robust computing capacity, care delivery can be more predictive, efficient, and effective in the diagnosis and care of an individual patient which improves outcomes for all parties: payers, providers, patients. As health systems look ahead toward affordable care and value-based payment systems, effective and efficient care of patients is tantamount to health and longevity of the organization and the right thing to do for the patient. But, that alone is not enough. To create change and to ultimately impact patients, these advances must make it "to the bedside" and truly impact the care of the patient at their most vulnerable time and in the most expensive setting in health care. Once these advances cross the chasm and change the way care is delivered inside the hospital (in a sustainable and lasting way for the patients, families, and providers) then it will be transformative and will truly change the hospital systems and health care for the better.

State Government

Responsible stewardship of taxpayer dollars is arguably the paramount duty of lawmakers throughout the United States who are tasked with using those dollars to maintain government operations while also ensuring that appropriations are made to achieve expectations of taxpayers. Achieving this balance is no simple matter. Texas in particular faces additional difficulties due to certain limitations, including a constitutional balanced budget requirement[98] and the adoption of a biennial budget which necessitates substantial estimations[99]. To aid with these constraints, the Texas Legislature often involves subject matter experts to educate lawmakers on how to best achieve desired policy outcomes with limited resources. However, this type of short-term partnership - where, for example, experts are only called upon to testify for a single committee hearing - has the potential to fall short as facts and data are either forgotten in the deluge of legislative issues or overlooked for political expediency. A variety of factors can cause this, but notably, the lack of a strong and well-established partnership between government and experts is a key contributor.

It is empirically proven in Texas that long-term partnerships, particularly with industry and academia, are both effective and fruitful. The Cancer Prevention and Research Institute of Texas (CPRIT), which was established by the Legislature and approved by Texas voters in 2007 to aid in cancer research and implement the Texas Cancer Plan, is a prime example. Since its inception, CPRIT reports recruiting 192 cancer researchers and labs, producing a Nobel Prize recipient, and awarding 1,452 grants

totaling $2.4 billion[100]. It has also provided the Legislature with evidence-based prevention interventions and services with sixty-six active projects that, combined, impacts every single one of the state's 254 counties[101]. This partnership has produced both desired health outcomes and significant cost-savings for the Legislature and Texans, which likely contributed to the Legislature's overwhelming support, and eventual voter approval, for Proposition 6 in 2019 to increase CPRIT's ability to award funding from a total of $3 billion to $6 billion[102]. Additionally, Texas is now seeking to replicate its success with CPRIT in the mental health arena with the recent passage of Senate Bill 11 by the 86th Legislature in 2019 to create the Texas Child Mental Health Care Consortium[103].

The benefits of a strong partnership are clear - improved research capabilities, healthier outcomes, and cost-efficient investments. As such, a partnership on computational medicine between the Texas government and other health care stakeholders has virtually limitless possibilities. Every area of medicine can benefit from a more individualized approach, and establishing the infrastructure for various computational models paves the path to achieving desired health outcomes, whether that outcome is lowering maternal mortality rates or simply increasing wellness visits to emphasize preventive care. Further, a computational medicine partnership could assist the Legislature with its ongoing efforts to contain health care costs in Texas, particularly in Medicaid. The most recent iteration of this initiative is delineated in the 86th Legislature's House Bill 1, the state's budget for fiscal years 2020 and 2021, and requires the Health and Human Services Commission (HHSC) realize at least $350 million in savings[104]. Currently, HHSC is limited to achieving these services by addressing systemic fraud, waste, and abuse as well as maximizing the use of federal Medicaid dollars. With computational medicine, the options expand as, for instance, expensive treatments are avoided with individualized decision making and preventive care, medications best for a patient are prescribed, and unnecessary procedures are cut back with individually best practices. Simply put, having the ability to take a pertinent health issue and produce cost-efficient solutions that will ensure a healthier population with striking accuracy would make the Legislature's job elementary - to the benefit of lawmakers and, more importantly, Texans.

Federal Government

Federal agencies can serve as enablers or even doers for issues with societal implications that are beyond visible market forces. There are elements of computational medicine that may require such actions from the federal government. The Human Genome Project, which evolved into a multi-agency, international and public-private partnership, is an example of a large scale effort that was driven by a federal recognition that technological convergence of advancements in robotics, image processing, data-base restructuring, computing, lasers and so forth, all outside of medicine, could be foundationally transformative to genomics. The scale of data that became to be generated from sequencing has been pivotal to driving data-centered analysis into the fabric of medicine. Learning from rich data sets remains a challenge as the questions being asked are more complex and the growth of available data

continues to stress the leading edge technologies, including artificial intelligence (AI) systems.

The promise of computational medicine resides in the complex technical landscape that spans more traditional supercomputing, through data sciences, cognitive computing and AI. The advancements in computational medicine will be furthered when we recognize which efforts are hindered by viewing the computational demands as a post-hoc add-on. Our notable successes in the application of computational science to decision making in high consequence situations have at their core teams of specialists from the outset that co-develop everything from the technologies to tools and share in the responsibilities of the outcomes. It is quite likely that AI based analog will need to develop in the same way. Many tough problems remain in prediction from models or data or both – problems that do not have visible economic drivers behind them. It is in this space that federal agencies can play a role in filling gaps that can help in making progress in the areas we have discussed. It is likely that at the federal level, that progress here will require two or more federal agencies participating. There are many means to organize such efforts, and finding the suitable champions is a key ingredient.

As enablers, the federal agencies can also help develop the technical base from academic programs to support the future workforce as well as key areas such as uncertainty quantification for AI, the areas of data trust and integrity as well as decision support. The United States has a unique ability today to shape this future. Defining the right partnerships and working together can transform health research and health care for us all (Figure 8).

Figure 8. The Texas Advanced Computing Center at The University of Texas at Austin is home to Frontera, the fastest supercomputer at any university and the 5th most powerful system in the world. This is the type of computational infrastructure that will allow breakthroughs in computational medicine.

REFERENCES

1. OSLER W, BEAN RB, BEAN WB. *Sir William Osler aphorisms : from his bedside teachings and writings.* Springfield, Ill.: Thomas; Number of pages.
2. CHENG YW, SNOWDEN JM, HANDLER SJ, TAGER IB, HUBBARD AE, CAUGHEY AB. Litigation in obstetrics: does defensive medicine contribute to increases in cesarean delivery? J Matern Fetal Neonatal Med 2014;27:1668-75.
3. WOLF JH. Risk and Reputation: Obstetricians, Cesareans, and Consent. J Hist Med Allied Sci 2018;73:7-28.
4. LIU P, SHI Z, WANG C, et al. The false-positive and false-negative predictive value of HIV antibody test in the Chinese population. J Med Screen 2008;15:72-5.
5. SCHORK NJ. Personalized medicine: Time for one-person trials. Nature 2015;520:609-11.
6. GIGERENZER G, GAISSMAIER W, KURZ-MILCKE E, SCHWARTZ LM, WOLOSHIN S. Helping Doctors and Patients Make Sense of Health Statistics. Psychol Sci Public Interest 2007;8:53-96.
7. WEGWARTH O, GIGERENZER G. The Barrier to Informed Choice in Cancer Screening: Statistical Illiteracy in Physicians and Patients. Recent Results Cancer Res 2018;210:207-21.
8. WEGWARTH O, GIGERENZER G. US gynecologists' estimates and beliefs regarding ovarian cancer screening's effectiveness 5 years after release of the PLCO evidence. Sci Rep 2018;8:17181.
9. GIGERENZER G, EDWARDS A. Simple tools for understanding risks: from innumeracy to insight. BMJ 2003;327:741-4.
10. FISCHHOFF B. Hindsight not equal to foresight: the effect of outcome knowledge on judgment under uncertainty. 1975. Qual Saf Health Care 2003;12:304-11; discussion 11-2.
11. TVERSKY A, KAHNEMAN D. Judgment under Uncertainty: Heuristics and Biases. Science 1974;185:1124-31.
12. JR AMCJJLJSGFJ. Overview of the 2015 American Congress of Obstetricians and Gynecologists' Survey on Professional Liability. 2015.
13. NELSON KB, DAMBROSIA JM, TING TY, GRETHER JK. Uncertain value of electronic fetal monitoring in predicting cerebral palsy. N Engl J Med 1996;334:613-8.
14. TURNER JM, MITCHELL MD, KUMAR SS. The physiology of intrapartum fetal compromise at term. Am J Obstet Gynecol 2020;222:17-26.
15. Executive summary: Neonatal brachial plexus palsy. Report of the American College of Obstetricians and Gynecologists' Task Force on Neonatal Brachial Plexus Palsy. Obstet Gynecol 2014;123:902-4.
16. GIBBINS KJ, WEBER T, HOLMGREN CM, PORTER TF, VARNER MW, MANUCK TA. Maternal and fetal morbidity associated with uterine rupture of the unscarred uterus. Am J Obstet Gynecol 2015;213:382 e1-6.
17. AL-ZIRQI I, DALTVEIT AK, VANGEN S. Infant outcome after complete uterine rupture. Am J Obstet Gynecol 2018;219:109 e1-09 e8.
18. HOYERT DL MOA. Maternal mortality in the United States: Changes in coding, publication, and data release, 2018. Hyattsville, MD: National Center for Health Statistics.: National Vital Statistics Reports, 2020 (vol 69).

19. KRAMER MR, STRAHAN AE, PRESLAR J, et al. Changing the conversation: applying a health equity framework to maternal mortality reviews. Am J Obstet Gynecol 2019;221:609 e1-09 e9.
20. DJULBEGOVIC B, GUYATT GH. Progress in evidence-based medicine: a quarter century on. Lancet 2017;390:415-23.
21. GREENHALGH T, HOWICK J, MASKREY N, EVIDENCE BASED MEDICINE RENAISSANCE G. Evidence based medicine: a movement in crisis? BMJ 2014;348:g3725.
22. EVIDENCE BC. Clinical Evidence. How much do we know?, 2013 (vol 2018).
23. MCALISTER FA, CAMPBELL NR, ZARNKE K, LEVINE M, GRAHAM ID. The management of hypertension in Canada: a review of current guidelines, their shortcomings and implications for the future. CMAJ 2001;164:517-22.
24. MCMURRAY J, SWEDBERG K. Treatment of chronic heart failure: a comparison between the major guidelines. Eur Heart J 2006;27:1773-7.
25. MCALISTER FA, VAN DIEPEN S, PADWAL RS, JOHNSON JA, MAJUMDAR SR. How evidence-based are the recommendations in evidence-based guidelines? PLoS Med 2007;4:e250.
26. STEEN RG, DAGER SR. Evaluating the evidence for evidence-based medicine: are randomized clinical trials less flawed than other forms of peer-reviewed medical research? FASEB J 2013;27:3430-6.
27. IOANNIDIS JP. Contradicted and initially stronger effects in highly cited clinical research. JAMA 2005;294:218-28.
28. STEGENGA J. Is meta-analysis the platinum standard of evidence? Stud Hist Philos Biol Biomed Sci 2011;42:497-507.
29. ROTHWELL PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet 2005;365:82-93.
30. GOLDSTEIN DB, NEED AC, SINGH R, SISODIYA SM. Potential genetic causes of heterogeneity of treatment effects. Am J Med 2007;120:S21-5.
31. GREENFIELD S, KRAVITZ R, DUAN N, KAPLAN SH. Heterogeneity of treatment effects: implications for guidelines, payment, and quality assessment. Am J Med 2007;120:S3-9.
32. KENT DM, HAYWARD RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. JAMA 2007;298:1209-12.
33. KENT DM, ROTHWELL PM, IOANNIDIS JP, ALTMAN DG, HAYWARD RA. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. Trials 2010;11:85.
34. ROTHWELL PM. Can overall results of clinical trials be applied to all patients? Lancet 1995;345:1616-9.
35. ROTHWELL PM, MEHTA Z, HOWARD SC, GUTNIKOV SA, WARLOW CP. Treating individuals 3: from subgroups to individuals: general principles and the example of carotid endarterectomy. Lancet 2005;365:256-65.
36. GLASZIOU PP, IRWIG LM. An evidence based approach to individualising treatment. BMJ 1995;311:1356-9.
37. IOANNIDIS JP, LAU J. The impact of high-risk patients on the results of clinical trials. J Clin Epidemiol 1997;50:1089-98.
38. HAYWARD RA, KENT DM, VIJAN S, HOFER TP. Reporting clinical trial results to inform providers, payers, and consumers. Health Aff (Millwood) 2005;24:1571-81.

39. KENT DM, RUTHAZER R, SELKER HP. Are some patients likely to benefit from recombinant tissue-type plasminogen activator for acute ischemic stroke even beyond 3 hours from symptom onset? Stroke 2003;34:464-7.
40. ROTHWELL PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. Lancet 2005;365:176-86.
41. SMEETH L, HAINES A, EBRAHIM S. Numbers needed to treat derived from meta-analyses--sometimes informative, usually misleading. BMJ 1999;318:1548-51.
42. KRAVITZ RL, DUAN N, BRASLOW J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. Milbank Q 2004;82:661-87.
43. KENT DM, ALSHEIKH-ALI A, HAYWARD RA. Competing risk and heterogeneity of treatment effect in clinical trials. Trials 2008;9:30.
44. KENNEDY-MARTIN T, CURTIS S, FARIES D, ROBINSON S, JOHNSTON J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. Trials 2015;16:495.
45. JAYASINGHE S. Complexity science to conceptualize health and disease: is it relevant to clinical medicine? Mayo Clin Proc 2012;87:314-9.
46. LIPSITZ LA. Understanding health care as a complex system: the foundation for unintended consequences. JAMA 2012;308:243-4.
47. MA'AYAN A. Complex systems biology. J R Soc Interface 2017;14.
48. PLSEK PE, GREENHALGH T. Complexity science: The challenge of complexity in health care. BMJ 2001;323:625-8.
49. COOK RI. How Complex Systems Fail, 2000.
50. SYMONS JB, F. How Computational Models Predict the Behavior of Complex Systems. Found Sci 2013;18:809–21.
51. AMERICAN COLLEGE OF O, GYNECOLOGISTS, TASK FORCE ON HYPERTENSION IN P. Hypertension in pregnancy. Report of the American College of Obstetricians and Gynecologists' Task Force on Hypertension in Pregnancy. Obstet Gynecol 2013;122:1122-31.
52. GALE C, MODI N, JAWAD S, et al. The WHEAT pilot trial-WithHolding Enteral feeds Around packed red cell Transfusion to prevent necrotising enterocolitis in preterm neonates: a multicentre, electronic patient record (EPR), randomised controlled point-of-care pilot trial. BMJ Open 2019;9:e033543.
53. SMITH GC, ROWITCH D, MOL BW. The role of prenatal steroids at 34-36 weeks of gestation. Arch Dis Child Fetal Neonatal Ed 2017;102:F284-F85.
54. CROWTHER CA, DODD JM, HILLER JE, HASLAM RR, ROBINSON JS, BIRTH AFTER CAESAREAN STUDY G. Planned vaginal birth or elective repeat caesarean: patient preference restricted cohort with nested randomised trial. PLoS Med 2012;9:e1001192.
55. SMITH GC, FLEMING KM, WHITE IR. Birth order of twins and risk of perinatal death related to delivery in England, Northern Ireland, and Wales, 1994-2003: retrospective cohort study. BMJ 2007;334:576.
56. BARRETT JF, HANNAH ME, HUTTON EK, et al. A randomized trial of planned cesarean or vaginal delivery for twin pregnancy. N Engl J Med 2013;369:1295-305.
57. SMITH GC, PELL JP, DOBBIE R. Birth order, gestational age, and risk of delivery related perinatal death in twins: retrospective cohort study. BMJ 2002;325:1004.

58. ZAFARMAND MH, GOOSSENS S, TAJIK P, et al. Planned Cesarean or planned vaginal delivery for twins: a secondary analysis of a randomized controlled trial. Ultrasound Obstet Gynecol 2019.
59. HORTON R. Offline: The gravy train of systematic reviews. Lancet 2019;394:1790.
60. PRIOR M, HIBBERD R, ASEMOTA N, THORNTON JG. Inadvertent P-hacking among trials and systematic reviews of the effect of progestogens in pregnancy? A systematic review and meta-analysis. BJOG 2017;124:1008-15.
61. BRAITHWAITE RS. A piece of my mind. EBM's six dangerous words. JAMA 2013;310:2149-50.
62. TOROUS J, KIANG MV, LORME J, ONNELA JP. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research. JMIR Ment Health 2016;3:e16.
63. CENTER PR. Mobile phone ownership over time, 2019.
64. SKINNER AL, ATTWOOD AS, BADDELEY R, EVANS-REEVES K, BAULD L, MUNAFO MR. Digital phenotyping and the development and delivery of health guidelines and behaviour change interventions. Addiction 2017;112:1281-85.
65. ACCENTURE. Accenture 2018 Consumer Survey on Digital Health, 2018.
66. DUNN J, RUNGE R, SNYDER M. Wearables and the medical revolution. Per Med 2018;15:429-48.
67. LARKIN A, HYSTAD P. Towards Personal Exposures: How Technology Is Changing Air Pollution and Health Research. Curr Environ Health Rep 2017;4:463-71.
68. REINERTSEN E, CLIFFORD GD. A review of physiological and behavioral monitoring with digital sensors for neuropsychiatric illnesses. Physiol Meas 2018;39:05TR01.
69. CHOUDHURY MD, COUNTS S, HORVITZ E. Predicting postpartum changes in emotion and behavior via social media*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Paris, France: Association for Computing Machinery, 2013.
70. VEGESNA A, TRAN M, ANGELACCIO M, ARCONA S. Remote Patient Monitoring via Non-Invasive Digital Technologies: A Systematic Review. Telemed J E Health 2017;23:3-17.
71. COVENEY PV, DOUGHERTY ER, HIGHFIELD RR. Big data need big theory too. Philos Trans A Math Phys Eng Sci 2016;374.
72. IYENGAR R, ALTMAN RB, TROYANSKYA O, FITZGERALD GA. MEDICINE. Personalization in practice. Science 2015;350:282-3.
73. POPPER KR. *The logic of scientific discovery.* New York,: Basic Books; Number of pages.
74. YANKEELOV TE, ATUEGWU N, HORMUTH D, et al. Clinically relevant modeling of tumor growth and treatment response. Sci Transl Med 2013;5:187ps9.
75. DOUGLAS PS, PONTONE G, HLATKY MA, et al. Clinical outcomes of fractional flow reserve by computed tomographic angiography-guided diagnostic strategies vs. usual care in patients with suspected coronary artery disease: the prospective longitudinal trial of FFR(CT): outcome and resource impacts study. Eur Heart J 2015;36:3359-67.
76. HOLMBERG L, VICKERS A. Evaluation of prediction models for decision-making: beyond calibration and discrimination. PLoS Med 2013;10:e1001491.
77. OLCHANSKI N, COHEN JT, NEUMANN PJ, WONG JB, KENT DM. Understanding the Value of Individualized Information: The Impact of Poor Calibration or Discrimination in Outcome Prediction Models. Med Decis Making 2017;37:790-801.
78. ALBA AC, AGORITSAS T, WALSH M, et al. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. JAMA 2017;318:1377-84.

79. WESSLER BS, LAI YH L, KRAMER W, et al. Clinical Prediction Models for Cardiovascular Disease: Tufts Predictive Analytics and Comparative Effectiveness Clinical Prediction Model Database. Circ Cardiovasc Qual Outcomes 2015;8:368-75.
80. ESCOBAR GJ, GUPTA NR, WALSH EM, SOLTESZ L, TERRY SM, KIPNIS P. Automated early detection of obstetric complications: theoretic and methodologic considerations. Am J Obstet Gynecol 2019;220:297-307.
81. RILEY RD, ENSOR J, SNELL KI, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. BMJ 2016;353:i3140.
82. MCKENNA MT, WEIS JA, BROCK A, QUARANTA V, YANKEELOV TE. Precision Medicine with Imprecise Therapy: Computational Modeling for Chemotherapy in Breast Cancer. Transl Oncol 2018;11:732-42.
83. STEYERBERG EW, VERGOUWE Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J 2014;35:1925-31.
84. VENKATESH KK, STRAUSS RA, GROTEGUT CA, et al. Machine Learning and Statistical Models to Predict Postpartum Hemorrhage. Obstet Gynecol 2020;135:935-44.
85. SOMMERS BD, ZECKHAUSER R. Probabilities and preferences: what economics can teach doctors and patients making difficult treatment decisions. Urol Oncol 2008;26:669-73.
86. BRASHERS DE. Communication and Uncertainty Management. Journal of Communication 2001;51:477-97.
87. KAHLOR L. PRISM: a planned risk information seeking model. Health Commun 2010;25:345-56.
88. AJZEN I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991;50:179-211.
89. AFIFI WA, WEINER JL. Toward a Theory of Motivated Information Management. Communication Theory 2006;14:167-90.
90. WITTE K. Putting the fear back into fear appeals: The extended parallel process model. Communication Monographs 1992;59:329-49.
91. FREIMUTH VS, STEIN JA, KEAN TJ. *Searching for health information : the Cancer Information Service model.* Philadelphia: University of Pennsylvania Press; Number of pages.
92. GRIFFIN RJ, DUNWOODY S, NEUWIRTH K. Proposed model of the relationship of risk information seeking and processing to the development of preventive behaviors. Environ Res 1999;80:S230-S45.
93. JOHNSON JD. *Cancer-related information seeking.* Hampton Press; Number of pages.
94. DOUGLAS M, WILDAVSKY AB. *Risk and culture : an essay on the selection of technical and environmental dangers.* Berkeley: University of California Press; Number of pages.
95. STEPHENS KK, ROBERTSON BW, MURTHY D. Throw me a lifeline: Articulating mobile social network dispersion and the social construction of risk in rescue communication. Mobile Media & Communication 2019:2050157919846522.
96. MCCOMAS KA. Defining moments in risk communication research: 1996-2005. J Health Commun 2006;11:75-91.
97. COMMITTEE ON OBSTETRIC P. Committee Opinion No. 697: Planned Home Birth. Obstet Gynecol 2017;129:e117-e22.
98. TEXAS C. Texas Constitution, Article III, Sec. 49a.
99. TEXAS C. Texas Constitution, Article III, Sec. 5.

100. TEXAS SO. Cancer Prevention & Research Institute of Texas, 2019.
101. TEXAS CPARIO. Populations Served by CPRIT Prevention Projects, 2019.
102. TEXAS SO. Proposition 6: Increasing CPRIT's bond authority from $3 billion to $6 billion, 2019.
103. TEXAS SO. Senate Bill 11: Texas Child Mental Health Care Consortium, 2019.
104. TEXAS SO. General appropriations act for the 2020-21 biennium, 2019.

FIGURE LEGENDS

Figure 1. Simplified idealized representation of the complex system, such as the human body or health care system. The system made of a vast number of highly interacting with each other components. Complex systems have, due to their structure, many different ways or causative pathways an outcome can occur, the phenomenon known as equifinality. Equifinality demands individualization in approach to the complex systems such as individualization of medical decision making.

Figure 2. The Computational Health Conference, held in Austin, TX in October of 2018. The conference brought together key stakeholders in health care and experts in computation and health from academia, government, industry, philanthropy, and communities with the goal of identifying future directions and opportunities for computational health and medicine.

Figure 3. Schematic representation of the “levels of evidence”. RCT denotes randomized controlled trial and SR denotes systematic review.

Figure 4. Proportion of U.S. adults who own cellphones and smartphones.

Figure 5. Monitoring interface of digital phenotype data streamed real-time from a patient’s cell phone to the supercomputer.

Figure 6. The role of computation in medicine. The human body and the health care system are complex systems, networks of highly coupled components intensely interacting with each other. These interactions give those systems redundancy, thus robustness to failure and, at the same time, equifinality, many different causative pathways leading to the same outcome. The equifinality demands individualization of medical care, which is urgently needed. Computational models excel in accounting for a very large number of interactions, thus in the modeling of complex systems, and hence enable in the individualization of medicine. They have the potential to enable individualization of medical decision making and, in consequence, better health outcomes and lower costs.

Figure 7. Individualized decision making in medicine can be thought of as composed of three consecutive parts: prediction of the individual probability of outcome; weighing of those individual probabilities of outcomes and individual patient preferences for outcomes; communication of the risk, the probability of the outcome and the burden of this outcome.

Figure 8. The Texas Advanced Computing Center at The University of Texas at Austin is home to Frontera, the fastest supercomputer at any university and the 5th most powerful system in the world. This is the type of computational infrastructure which will allow breakthroughs in the computational medicine.