EDM Forum
EDM Forum Community

eGEMs (Generating Evidence & Methods to improve patient outcomes) | Publish

3-7-2016

Consensus Statement on Electronic Health Predictive Analytics: A Guiding Framework to Address Challenges

Ruben Amarasingham
PCCI, ruben.amarasingham@phhs.org

Anne-Marie J. Audet
The Commonwealth Fund, amjaudet@post.harvard.edu

David W. Bates
Brigham and Women's Hospital, dbates@partners.org

I. Glenn Cohen
Harvard Law School, igcohen@law.harvard.edu

See next pages for additional authors

Follow this and additional works at: http://repository.edm-forum.org/egems

Part of the Health Information Technology Commons, Health Services Research Commons, and the Other Public Health Commons

Recommended Citation

Amarasingham, Ruben; Audet, Anne-Marie J.; Bates, David W.; Cohen, I. Glenn; Entwistle, Martin; Escobar, G J.; Liu, Vincent; Etheredge, Lynn; Lo, Bernard; Ohno-Machado, Lucila; Ram, Sudha; Saria, Suchi; Schilling, Lisa M.; Shah, Anand; Stewart, Walter F.; Steyerberg, Ewout W.; and Xie, Bin (2016) "Consensus Statement on Electronic Health Predictive Analytics: A Guiding Framework to Address Challenges," *eGEMs (Generating Evidence & Methods to improve patient outcomes)*: Vol. 4: Iss. 1, Article 3.
DOI: http://dx.doi.org/10.13063/2327-9214.1163
Available at: http://repository.edm-forum.org/egems/vol4/iss1/3

This Informatics Model/Framework is brought to you for free and open access by the the Publish at EDM Forum Community. It has been peer-reviewed and accepted for publication in eGEMs (Generating Evidence & Methods to improve patient outcomes).

The Electronic Data Methods (EDM) Forum is supported by the Agency for Healthcare Research and Quality (AHRQ), Grant 1U18HS022789-01. eGEMs publications do not reflect the official views of AHRQ or the United States Department of Health and Human Services.

Consensus Statement on Electronic Health Predictive Analytics: A Guiding Framework to Address Challenges

Abstract

Context: The recent explosion in available electronic health record (EHR) data is motivating a rapid expansion of electronic health care predictive analytic (e-HPA) applications, defined as the use of electronic algorithms that forecast clinical events in real time with the intent to improve patient outcomes and reduce costs. There is an urgent need for a systematic framework to guide the development and application of e-HPA to ensure that the field develops in a scientifically sound, ethical, and efficient manner.

Objectives: Building upon earlier frameworks of model development and utilization, we identify the emerging opportunities and challenges of e-HPA, propose a framework that enables us to realize these opportunities, address these challenges, and motivate e-HPA stakeholders to both adopt and continuously refine the framework as the applications of e-HPA emerge.

Methods: To achieve these objectives, 17 experts with diverse expertise including methodology, ethics, legal, regulation, and health care delivery systems were assembled to identify emerging opportunities and challenges of e-HPA and to propose a framework to guide the development and application of e-HPA.

Findings: The framework proposed by the panel includes three key domains where e-HPA differs qualitatively from earlier generations of models and algorithms (Data Barriers, Transparency, and Ethics) and areas where current frameworks are insufficient to address the emerging opportunities and challenges of e-HPA (Regulation and Certification; and Education and Training). The following list of recommendations summarizes the key points of the framework:

1. Data Barriers: Establish mechanisms within the scientific community to support data sharing for predictive model development and testing.
2. Transparency: Set standards around e-HPA validation based on principles of scientific transparency and reproducibility.
3. Ethics: Develop both individual-centered and society-centered risk-benefit approaches to evaluate e-HPA.
4. Regulation and Certification: Construct a self-regulation and certification framework within e-HPA.
5. Education and Training: Make significant changes to medical, nursing, and paraprofessional curricula by including training for understanding, evaluating, and utilizing predictive models.

Acknowledgements

GRANT SUPPORT: Gordon and Betty Moore Foundation Framework and Action Plan for Predictive Analytics Grant #3861 Acknowledgements: The authors wish to thank Adeola Jaiyeola, MD, MHSc, Nam Nguyen, BA, and other individuals of PCCI for their contributions to the preparation of this manuscript.

Keywords

Informatics, Health Information Technology, Ethics, Clinical decision support systems, electronic predictive analytics, predictive models, big data

Disciplines

Health Information Technology | Health Services Research | Other Public Health

This model/framework is available at EDM Forum Community: http://repository.edm-forum.org/egems/vol4/iss1/3

Creative Commons License

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.

Authors

Ruben Amarasingham, *PCCI*; Anne-Marie J Audet, *The Commonwealth Fund*; David W Bates, *Brigham and Women's Hospital*; I. Glenn Cohen, *Harvard Law School*; Martin Entwistle, *Palo Alto Medical Foundation*; G J Escobar; Vincent Liu, *Kaiser Permanente*; Lynn Etheredge; Bernard Lo, *The Greenwall Foundation*; Lucila Ohno-Machado, *University of California, San Diego*; Sudha Ram, *Management Information Systems, University of Arizona, Tucson, AZ, United States.*; Suchi Saria, *Johns Hopkins University*; Lisa M Schilling, *University of Colorado, School of Medicine*; Anand Shah, *PCCI*; Walter F Stewart, *Sutter Health*; Ewout W Steyerberg, *Erasmus MC*; Bin Xie, *PCCI*.

This model/framework is available at EDM Forum Community: http://repository.edm-forum.org/egems/vol4/iss1/3

eGEMs
Generating Evidence & Methods to improve patient outcomes

Consensus Statement on Electronic Health Predictive Analytics: A Guiding Framework to Address Challenges

Ruben Amarasingham, MD;[i] Anne-Marie J. Audet, MD;[ii] David W. Bates, MD, MSc;[iii] I. Glenn Cohen;[iv] Martin Entwistle;[v] G. J. Escobar; Vincent Liu, MD, MS;[vi] Lynn Etheredge; Bernard Lo, MD;[vii], Lucila Ohno-Machado, MD, PhD;[viii] Sudha Ram, PhD;[ix] Suchi Saria, PhD;[x] Lisa M. Schilling, MD, MSPH;[xi] Anand Shahi, CFA, CFP;[i] Walter F. Stewart, PhD;[xii] Ewout W. Steyerberg, PhD, MSc;[xiii] Bin Xie

ABSTRACT

Context: The recent explosion in available electronic health record (EHR) data is motivating a rapid expansion of electronic health care predictive analytic (e-HPA) applications, defined as the use of electronic algorithms that forecast clinical events in real time with the intent to improve patient outcomes and reduce costs. There is an urgent need for a systematic framework to guide the development and application of e-HPA to ensure that the field develops in a scientifically sound, ethical, and efficient manner.

Objectives: Building upon earlier frameworks of model development and utilization, we identify the emerging opportunities and challenges of e-HPA, propose a framework that enables us to realize these opportunities, address these challenges, and motivate e-HPA stakeholders to both adopt and continuously refine the framework as the applications of e-HPA emerge.

Methods: To achieve these objectives, 17 experts with diverse expertise including methodology, ethics, legal, regulation, and health care delivery systems were assembled to identify emerging opportunities and challenges of e-HPA and to propose a framework to guide the development and application of e-HPA.

[i]PCCI, [ii]The Commonwealth Fund, [iii]Brigham and Women's Hospital, [iv]Harvard Law School, [v]Palo Alto Medical Foundation, [vi]Kaiser Permanente, [vii]The Greenwall Foundation, [viii]University of California, San Diego, [ix]Management Information Systems, University of Arizona, [x]Johns Hopkins University, [xi]University of Colorado, School of Medicine, [xii]Sutter Health, [xiii]Erasmus Medical Center

CONTINUED

Findings: The framework proposed by the panel includes three key domains where e-HPA differs qualitatively from earlier generations of models and algorithms (Data Barriers, Transparency, and Ethics) and areas where current frameworks are insufficient to address the emerging opportunities and challenges of e-HPA (Regulation and Certification; and Education and Training). The following list of recommendations summarizes the key points of the framework:

1. Data Barriers: Establish mechanisms within the scientific community to support data sharing for predictive model development and testing.
2. Transparency: Set standards around e-HPA validation based on principles of scientific transparency and reproducibility.
3. Ethics: Develop both individual-centered and society-centered risk-benefit approaches to evaluate e-HPA.
4. Regulation and Certification: Construct a self-regulation and certification framework within e-HPA.
5. Education and Training: Make significant changes to medical, nursing, and paraprofessional curricula by including training for understanding, evaluating, and utilizing predictive models.

Background

Algorithms and models have long been used in health care to assist decision-making. There are well-established frameworks that provide guidance to the development and utilization of these algorithms and models. A well-known example of an algorithm is the Acute Physiology and Chronic Health Examination (APACHE) II score for intensive care unit mortality.[1] This model uses a few variables that can be calculated by hand without a computer. A well-known framework that provides guidance for the development and utilization of such algorithms and models is the one proposed jointly by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) and the Society for Medical Decision Making (SMDM) through their ISPOR-SMDM modeling good research practices task force.[2-6]

e-HPA shares common features with earlier generation algorithms, models, and frameworks providing guidance to the initial development and implementation of e-HPA. However, the current quantity of input data and the growing sophistication of the algorithms and models in e-HPA are on a different scale than the earlier algorithms and models. The implementation of e-HPA on a wide scale to aid in real-time, point-of-care decision-making brings a new set of challenges and opportunities that are not covered by earlier frameworks. As a result of rapid development, many ethical, legal, regulatory, methodological, and technical challenges are emerging; consequently, the existing frameworks in these areas are not well equipped to provide sufficient guidance for addressing these new challenges.[7,8] Thus, there is an urgent need to identify these opportunities and

Volume 4 (2016) Issue Number 1

challenges in order to establish a framework to ensure robust development of the field.

To achieve these goals, the Gordon and Betty Moore Foundation (GBMF) provided a grant to Parkland Center for Clinical Innovation (PCCI), a nonprofit research and development organization in Dallas, TX, to assemble a panel of experts with a broad range of expertise to provide a framework for identifying and addressing the opportunities and challenges of e-HPA.

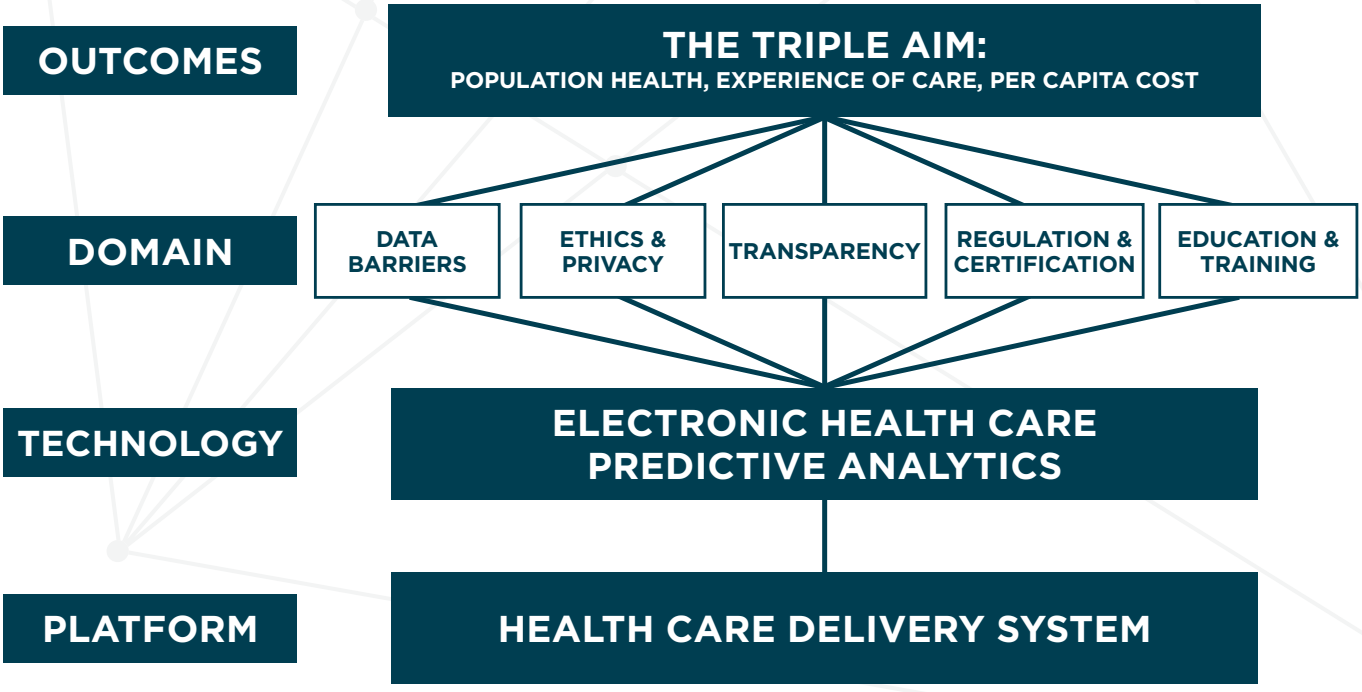
This paper reports the emerging opportunities and challenges identified and the framework (Figure 1) proposed by the panel. The proposed framework builds upon previous frameworks of model development and utilization with a focus on the new opportunities and challenges that are emerging in the era of big data. This framework can serve multiple purposes including, but not limited to the following:

1. Identifying knowledge gaps in the clinical application of predictive analytics and encouraging the research community to focus on filling these gaps;
2. Informing funding agencies to provide targeted funding opportunities for specific areas of e-HPA research that align with their missions; and
3. Initiating a discussion on how to establish the appropriate legal, regulatory, and ethical frameworks for health care organizations, governmental agencies, and other oversight agencies to make use of e-HPA.

Methodology

The panel members were selected by GBMF and PCCI staff to ensure broad representation of the field. The final panel includes 17 nationally or internationally recognized experts: from academia (7), private foundations (2), health care

Figure 1. Guiding Framework



Transparency: Why should I trust my care to a computer model?
Regulation & Certification: What's the right balance between regulation and certification on the one hand, and innovation on the other?
Data Barriers: Can modeling techniques keep pace with the size and complexity of data?
Ethics & Privacy: What should be the right framework to think through some of the ethics issues arising from e-HPA?
Education & Training: Can doctors keep pace with e-HPA?

delivery systems (3), model developers (3), and governmental agencies (2). Their collective expertise included data information technology, methodology, legal, regulatory, ethics, and health care delivery. Prior to the symposium, GBMF and PCCI staff carried out semistructured, in-depth interviews based on a topic guide with each individual panel member. This allowed for a detailed exploration to identify the key emerging opportunities and challenges, and to finalize the in-person symposium discussion topics. Interviews were audio recorded and transcribed verbatim with participant permission. A variation of content analysis was then used to develop a coding scheme for performing a qualitative description of the themes discussed by interviewees. The final codebook included both inductive and deductive codes that were finalized after reaching a consensus among the three-member research team at PCCI. As a result, the codebook was used to create a draft framework with three interrelated topics where e HPA differs qualitatively from earlier generations of models and algorithms (Data Barriers; Transparency; and Ethics). Additionally, the framework includes two areas where current supporting infrastructure frameworks are insufficient to address the emerging opportunities and challenges of e-HPA (Regulation and Legal; and Education and Training).

These five focus areas provide a structure for the framework. The PCCI research team's proposal to build the framework around these five areas was accepted unanimously by the panel. After these individual interviews, a PCCI writing team summarized an initial list of key points for each focus area and circulated it to all panelists prior to the symposium. The symposium began with a session that clarified the definition of e-HPA for all participants, reviewed the purposes and goals of the symposium, and reviewed the initial list of key points for each focus area. Panelists were then divided into five subgroups, each covering a focus area. These subgroups provided recommendations to realize the opportunities and address the challenges of each focus area. After the subgroup sessions, full panel collaboration synthesized the subgroup findings. The research team then reviewed all meeting transcripts and drafted a manuscript. All panelists provided feedback on earlier versions of the manuscript, and both reviewed and approved the final version of the manuscript.

Findings and Recommendations

Data Barriers and Model Development

Recommendation 1: Owners and potential users of high quality, de-identified, actionable, and diverse data sources should develop and implement a mutually agreed upon mechanism so that these data sources can be made available to the scientific community for development and testing of predictive models. Such mechanisms need proper governance as well as formal processes for ensuring collaboration in the development, testing, and validation of predictive models. Appropriate protocols to address data breaches or re-identification of de-identified data are also needed as even a robust de-identification process does not make re-identification impossible.[24]

High quality and diverse data sets fuel predictive models.[9,10] Historically, data used in models have been limited to the data collected during the patient encounter with primary access to this data from hospital EHRs, and laboratory, pharmacy, and administrative data sets.[11-13]

Comprehensive electronic record and other data sources, such as laboratory, pharmacy, and radiological reports that capture detailed patient data have now attained mainstream use in the United States.[14-16] For example, over 9 in 10 (93 percent) hospitals possessed a certified EHR technology in 2013, increasing by 29 percent since

Volume 4 (2016) Issue Number 1

2011.[17] In addition to making clinical data routinely available, electronic data sources also have greater vertical and horizontal connectivity (i.e., within and across health care institutions). One example of horizontal connectivity is data exchanges among health care providers via health information exchanges (HIEs). Having an e-HPA framework will provide guidance on the development, implementation, operation, and utilization of such HIEs. New interface capabilities, which allow patients to submit patient generated data via cell phones, personal computers, and personal monitoring devices, provide additional data sources that are collected externally to the health care environment.[18-21]

Creating new interface capabilities represents a major challenge—and opportunity—in medical informatics and health care management more broadly. For example, patient generated data, such as frequent blood sugar measurements recorded by user devices and transmitted to the health care providers, can provide much richer information than simpler data of such measurements taken during clinic visits, enabling a more timely and accurate diagnosis and treatment.

On the other hand, development of diverse and comprehensive data sets presents significant challenges and would require a concerted effort by stakeholders. For example, collection of patient-reported data requires engagement and commitment from patients. Similarly, obtaining data from social media, health forum websites, and even at-home monitoring technologies, involves coordination with IT developers as well as other stakeholders.

Patient level linkage of a wider array of disparate data sets allows for greater opportunities for data to be combined in a way that could threaten patient privacy.[22] Data collection methods from various sources would need to be standardized to improve protection of patient information. As a result, it is crucial to ensure the quality of linked data; usually, data from different sources need to be harmonized, cleaned, and reliably linked before e-HPA models can produce reliable results. For example, patient weight and height need to be recorded in the same units. When data variables are harmonized, such as is the case in many HIEs, providers in these exchanges can exchange information.

To realize such opportunities and address these challenges, the panel believes that expanding the scientific community's access and ability to share diverse data types with appropriate protections and use agreements could propel an era of impactful predictive modeling.

In addition to making patient-level clinical data accessible to the scientific community, there is a need to explore data capture opportunities in other settings. We need to look beyond administrative, clinical, and physiological variables, and capture and incorporate social, behavioral, patient-reported, patient-monitored, genetic, public health, and environmental data into predictive models.[14, 22-24] This could improve the accuracy and scope of predictive models.

Transparency and Model Evaluation

Recommendation 2: Developers of health care predictive analytics should adopt the principle of transparency and establish a consensus process involving key stakeholders to develop standards around validation and transparency of predictive modeling that include the following: (1) defining best modeling practices; (2) ensuring clinical coherence; (3) incorporating diverse data sources; (4) specifying guidelines for validation and unbiased evaluation; and (5) explaining model performance in a language clearly understandable by patients and clinicians.

Besides increased size and complexity of data, modeling techniques that can use such increasingly big and complex data to provide accurate predictions also present potential opportunities and challenges.

With increasing availability of data and computational power, models that make specific predictions around the effectiveness of alternative therapeutic plans for a specific patient may soon be routine. For example, a current heart failure program based on a crude prediction profile reserves the most intensive treatments for patients at highest risk and most likely to benefit,[26] but in a new paradigm, a predictive model could recommend specific clinical and care management strategies, such as specific drugs and follow-up periods, and even whether case management would be effective. Moreover, with the growing (and potential) use of genetics information for diagnostics and therapeutics, there are expanding opportunities for personalized medicine that will rely on e-HPA methods.[27]

On the other hand, these increasingly complex models may make transparency, defined here as "the practice of making key aspects of the models and algorithms available to enable review and unbiased replication of development and testing methods," more difficult. Moreover, independent evaluations of these models also become more challenging as new modeling technologies are becoming less understandable to nonexperts. For instance, a random forest- or neural-network method using hundreds of variables may be more difficult to interpret than a logistic regression method using 12 variables.[28] Publishing an unvalidated model or a model used inappropriately by users may lead to suboptimal results, resulting in skepticism or even backlash against all models—potentially hindering the development of e-HPA.

As the volume and level of detail of patient and other data increase, consideration must be given to the development of additional methodologies that—at first sight—are not directly related to predictive modeling. The major factor behind this need is the problem of bias. Suppose that a health system makes full-scale genomic testing available for a minimal co-pay and that X percent of its members sign up for it. A predictive model for condition Z developed from health system Y data could reach very wrong conclusions (e.g., reverse causation) due to the fact that people who volunteer for such testing may be very different from those who do not choose to volunteer. Other biases could creep in due to nonrandom distribution of incentives. Therefore, encouragement should be given to methodological research in this area. One broad methodological endeavor that should receive attention is that of estimation bias.[29] In addition, an existing methodological current—the use of counterfactual methods for comparative effectiveness research, which has a long history[30]—should continue to be encouraged, but with big data in mind.

To realize these opportunities and address the challenges, we recommend that the information be presented in a way that can be understood and applied by a wide range of users—including clinical informatics experts, providers, and patients. Transparency can increase provider and patient trust in the application of predictive analytics in health care.[31,32] Transparency should promote a more active collaborative scientific and vendor community and should allow those that apply e-PHA to select the best models based on their environment and available data. Transparency in health care predictive analytics must be carefully implemented. Instead of a one-size-fits-all approach, a transparency framework should be adaptable to tailoring a predictive model's prototype, complexity, and users.[31]

Volume 4 (2016) Issue Number 1

As part of the transparency framework, standards for the evaluation of predictive models need to be set.[33] Due to a lack of regulatory framework and the rapid development of health care predictive analytics, the quality of available models varies widely.[34-36] Published, peer-reviewed models provide a certain level of quality; however, the quality and extent of assessment of commercially available models is unclear.[44] There is a need to specify standards around validation and transparency of predictive modeling to ensure that models are robust and of high quality.[33,37]

A possible strategy for standardization is to establish best practice guidelines to define and help incentivize best modeling practices, clinical coherence (consistency with clinical practice), standards for incorporating diverse data sources, and appropriate frameworks for validation as well as unbiased evaluation. An example of a transparency reporting guideline is the Consolidated Standards of Reporting Trials (CONSORT) for reporting clinical trial methods.[38]

We suggest that such standards be developed in consultation using a consensus process that includes the end users of predictive modeling including researchers, clinical informatics experts, clinicians, and patients. In particular, potential end users need nontechnical, "plain English," information regarding the applicable population and performance characteristics of models. For example, rather than just reporting a single statistic (e.g., the c statistic),[39] end users need to know how many patients would have been evaluated and correctly identified at different thresholds in their specific clinical setting. Such information should be made available to these end users to the maximum extent allowable under the current framework of intellectual property protection.

Ethics

Recommendation 3: Users of health care predictive analytics should develop a risk-benefit analysis approach for the use of health care predictive analytics at the individual, organizational, and societal levels to determine the adoption of these models.

Besides opportunities and challenges presented by the increasingly complex data and modeling techniques, the increasing sophisticated utilization of e-HPA models in assisting medical decision-making also presents new opportunities and challenges. For example, e-HPA has been used to identify when care can be escalated or de escalated without significantly compromising care outcomes, allowing for optimal resource allocation. In one example of such utilization, patients predicted to have a high risk of readmission may be targeted to receive more intense care management. So, resources are allocated to the patients that are most likely to benefit.[26] While such applications of e-HPA are still emerging, the potential opportunities and benefits are huge and have been discussed elsewhere.[26]

On the other hand, using e-HPA to inform medical decision-making at the patient- and system levels raises important ethical issues that need to be addressed, especially when such models are used to escalate or de-escalate care for patients.

It is the responsibility of both the individual physician who uses the predictive models and the health care organization that embeds predictive models into their EHRs to ensure that the benefits of using the models outweigh the risks.[33] To do so, they need to rigorously assess the benefits and risks of implemented e-HPA in various situations, which includes the actions recommended or triggered by the model's results. Moreover, the distribution of

benefits and risks across the variety of patients to whom the e-PHA is applied needs to be analyzed. It is possible that the model may benefit the population as a whole but be seriously harmful to some subgroups or individuals.

Additionally, developers of EHR with built-in predictive analytics should explicitly communicate the risks and benefits to their clients. Clinical providers and patients using predictive analytics should carefully consider risks and benefits of using or not using predictive models prior to their use. It is possible that, in the future, an institution foregoing the use of predictive models or a physician failing to appropriately use predictive models could be considered outside of practice norms when there is clear evidence that the use of e-HPA in such situations is beneficial.

With regard to consent, for the clinical applications of the predictive analytics (PA) model, the issues are whether use of the PA should be considered a condition of care, similar to the use of a clinical decision support tool, and what type of notification of PA should be given to patients.

Regulation and Certification

Recommendation 4: e-HPA Practitioners should implement a carefully constructed self-regulation and certification framework for e-HPA. The goal of the framework would be to realize the potential of e-HPA by addressing a multitude of concerns—such as the assessment of benefits and risks, transparency, and patient privacy—without stifling innovation. Voluntary organizations may help fulfill this need.

Given these complex issues in data, transparency, and model evaluation, and ethics, establishing the appropriate regulation and certification framework for e-HPA becomes crucial.

Health care predictive analytics offers an unprecedented opportunity and ability to leverage massive amounts of data to improve the health of patients and lower the cost of health care.[9,14,40] This opportunity is not, however, without significant risks to patients.[15,31] Underperformance of models and negligent operation of predictive analytics by health care organizations or clinicians, for example, are conceivable scenarios in an unregulated market of predictive models.[41,42]

Given the nascent state of the e-HPA era, the need to balance oversight without stifling innovation cannot be overstated. A strict regulatory framework for e-HPA may discourage developers and research institutions from investigating and developing novel models for use in various health care sectors. Instead, we recommend as a first step, the need for a carefully constructed self-regulation and certification framework within e-HPA.[31,43] We define self-regulation as a process steered by key stakeholders, including health care researchers, predictive modelers, mathematicians, IT developers, manufacturers, clinicians, hospital executives, patients and their advocates, social workers and special societies aimed at minimizing patient risk, developing best practices, and setting standards for improving quality of care. Self-regulation may be applied in conjunction with some form of government regulation, or it could fill the vacuum in the absence of government oversight and regulation. Examples of voluntary organizations that may support self-regulation for e-HPA include the Joint Commission on Accreditation, Healthcare and Certification (JCAHO),[44] National Association for Healthcare Quality (NAHQ),[67] National Quality Forum (NQF),[45] and American Medical Informatics Association (AMIA).[46]

Certification standards based on model evaluation and documentation of performance metrics can further mitigate risks imposed on patients.[31] Model

evaluation and expanded documentation of a broader range of performance metrics could ensure quality application of predictive models in hospitals and other care settings.[33,47,48] Model evaluation should include stress testing of models with respect to differing populations, missing and erroneous data, and clarity regarding validation procedures.[33,37,49,50] One strategy to achieve this would be to have a third-party provider certify and evaluate models before their use in health care settings.[33] Certification would require that the models achieve and maintain certain performance metrics.

Education and Training

Recommendation 5: Medical schools and other professional training institutions should incorporate e-HPA competency into medical, nursing, and paraprofessional curricula and training.

The growing sophistication and performance of predictive models in health care will create new training requirements of the medical professionals.[51] An interdisciplinary workforce capable of evaluating and utilizing predictive models will be an essential ingredient in advancing the value of health care predictive analytics.[52] Specifically, individuals with expertise in both health care and quantitative modeling techniques and informatics are already highly desired.[53] There is also a need to attract individuals with sophisticated mathematical and modeling skill sets into the health care industry. Moreover, expertise from areas operating at the periphery of health care such as in biodesign, biomedical engineering, user experience and design, human factors engineering, and information security will be equally critical.

The panel voiced concern that the current workforce may not have the requisite professional training for an emerging predictive modeling paradigm in health care. Concerns range from clinicians' and administrators' poor understanding of concepts in probability, statistics, and heuristics[54] to medical, nursing, and paraprofessional curricula and training that do not emphasize the value of accurate data collection and potential for reuse.[52] An interdisciplinary trained workforce with a greater understanding of probability and statistics, not just predictive modeling techniques, could alleviate these concerns.[55] This would require medical, nursing, and paraprofessional curricula and training to adopt e-HPA as a competency.

Conclusion

As with earlier generations of models and algorithms, e-HPA at the point of care has the potential to lead to significant improvement in the quality, efficiency, and convenience of health care. However, to realize such potential, challenges exist in key technical areas including increasingly big and complex data sources, increasingly complex modeling techniques, increasingly sophisticated ways of using e-HPA to guide medical decision-making, and key supporting infrastructures—including regulation and certification, and education and training. While earlier generations of frameworks are still instructive, the opportunities and challenges are sufficiently different that a new framework is urgently needed to provide guidance to the nascent field.

In this paper, a panel with 17 nationally and internationally known experts in diverse fields proposed a framework that includes developing an oversight strategy and standards for transparency, privacy, risk-benefit analysis, data availability, and health care workforce training and education. While the field of e-HPA is still in its early stages and the framework presented will grow and adapt as the field evolves, the panel believes that the framework presented provided the scaffolding and a starting point upon which to build future development of the field. We call upon the field to update this framework

as the field evolves and new opportunities and challenges emerge.

Given the scope of topics covered and the limited space, we have not explored the implications or ramifications of these recommendations in detail. An in-depth discussions of these topics can be found elsewhere.[7,8,56] It is the task of health care leaders, e-HPA practitioners, and other stakeholders to ensure an infrastructure that ultimately promotes effective use of predictive analytics to improve patient outcomes, satisfaction, and the value of health care resources.

References

1. Knaus, W.A., et al., APACHE II: a severity of disease classification system. Crit Care Med, 1985. 13(10): p. 818-29.
2. Roberts M, Russell LB, Paltiel AD, et al. Conceptualizing a model: A report of the ISPOR-SMDM modeling good research practices task force-2. Value Health 2012;15:804-11.
3. Caro JJ, Briggs AH, Siebert U, et al. Modeling good research practices - overview: A report of the ISPOR-SMDM modeling good research practices task force-1. Value Health 2012;15:796-803.
4. Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: A report of the ISPOR-SMDM modeling good research practices task force-4. Value Health 2012;15:843-50.
5. Briggs AH, Weinstein MC, Fenwick E, et al. Model parameter estimation and uncertainty analysis: A report of the ISPOR-SMDM modeling good research practices task force-6. Value Health 2012;15:835-42.
6. Floyd J. Fowler,F.J. Shared Decision Making & Medical Costs. Foundation Perspectives. Informed Medical Decisions Foundation. 2012. Available at: http://www.informedmedicaldecisions.org/sdmmonth/wp-content/uploads/2014/02/Perspectives_SDM-Cost_v2.pdf
7. Cohen IG, A.R., Shah A, Xie B, Lo B., *The Legal And Ethical Concerns That Arise From Using Complex Predictive Analytics In Health Care*. Accepted for Publication in Health Affairs, 2014.
8. Amarasingham R, P.R., Huesch M, Nguyen NQ, Xie B., *Implementing Electronic Health Care Predictive Analytics: Considerations And Challenges*. Accepted for Publication in Health Affairs, 2014.
9. Edelstein, P., *Emerging Directions in Analytics*. Health Manag Technol, January, 2013.
10. Jha, A.K., et al., *Use of electronic health records in U.S. hospitals*. N Engl J Med, 2009. 360(16): p. 1628-38.
11. Cholleti, S., et al., *Leveraging derived data elements in data analytic models for understanding and predicting hospital readmissions*. AMIA Annu Symp Proc, 2012. 2012: p. 103-11.
12. Wang, L., et al., *Predicting risk of hospitalization or death among patients with heart failure in the veterans health administration*. Am J Cardiol, 2012. 110(9): p. 1342-9.
13. Amarasingham, R., et al., *An automated model to identify heart failure patients at risk for 30-day readmission or death using electronic medical record data*. Med Care, 2010. 48(11): p. 981-8.
14. Choi, I.Y., et al., *Perspectives on Clinical Informatics: Integrating Large-Scale Clinical, Genomic, and Health Information for Clinical Care*. Genomics Inform, 2013. 11(4): p. 186-190.
15. Rothman, B., J.C. Leonard, and M.M. Vigoda, *Future of electronic health records: implications for decision support*. Mt Sinai J Med, 2012. 79(6): p. 757-68.
16. Miriovsky, B.J., L.N. Shulman, and A.P. Abernethy, *Importance of health information technology, electronic health records, and continuously aggregating data to comparative effectiveness research and learning health care*. J Clin Oncol, 2012. 30(34): p. 4243-8.
17. Charles, D.; Gabriel, M.; Furukawa,M.F.Adoption of Electronic Health Record Systems among U.S. Non-federal Acute Care Hospitals: 2008-2013. ONC Data Brief, No. 16, May 2014.
18. Zhang , M., Velasco, F.T., Musser, R.C., Kawamoto K., *Enabling Cross-Platform Clinical Decision Support through Web-Based Decision Support in Commercial Electronic Health Record Systems: Proposal and Evaluation of Initial Prototype Implementations*. AMIA Annu Symp Proc, 2013. 2013: p. 1558–1567.
19. Murdoch, T.B., Detsky, A. S., *The inevitable application of big data to health care*. JAMA, 2013 309(13): p. 1351 - 1352.
20. Witonsky, P., *Leveraging EHR investments through medical device connectivity*. Healthc Financ Manage, 2012. 66(8): p. 50-3.
21. Pringle, S. and A. Lippitt, *Interoperability of electronic health records and personal health records: key interoperability issues associated with information exchange*. J Healthc Inf Manag, 2009. 23(3): p. 31-7.
22. Moore, G., et al., *Prospective validation of a predictive model that identifies homeless people at risk of re-presentation to the emergency department*. Australas Emerg Nurs J, 2012. 15(1): p. 2-13.
23. Shublaq, N.W. and P.V. Coveney, *Merging genomic and phenomic data for research and clinical impact*. Stud Health Technol Inform, 2012. 174: p. 111-5.
24. Chen, C., Weider, K., Konopka, K., Danid, M., *Incorporation of Socioeconomic Status Indicators into Policies for the Meaningful Use of Electronic Health Records*. Journal of Health Care for the Poor and Underserved February, 2014. 25(1): p. 1-16.
25. Li, F., et al., *New threats to health data privacy*. BMC Bioinformatics, 2011. 12 Suppl 12: p. S7.
26. Amarasingham, R., et al., *Allocating scarce resources in real-time to reduce heart failure readmissions: a prospective, controlled study*. BMJ Qual Saf, 2013. 22(12): p. 998-1005.
27. Smith, W.P., et al., *Validation of a modified Medical Resource Model for mass gatherings*. Prehosp Disaster Med, 2013. 28(1): p. 16-22.

28. Li, X. and Y. Wang, *Applying various algorithms for species distribution modelling*. Integr Zool, 2013. 8(2): p. 124-35.
29. Lash TL, Fox MP, Fink AK. Applying quantitative bias analysis to epidemiologic data. Springer, New York, 2009.
30. Rosenbaum, P. R. and D. B. Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects." Biometrika 70(1): 41-55.
31. Bates, D.W., *FDASIA Committee Report*. 2013: http://www.healthit.gov/facas/sites/faca/files/FDASIARecommendationsDraft030913_v2.pdf (Accessed March, 2014).
32. in *Health IT and Patient Safety: Building Safer Systems for Better Care*. 2011: Washington (DC).
33. Moons, K.G., et al., *Risk prediction models: II. External validation, model updating, and impact assessment*. Heart, 2012. 98(9): p. 691-8.
34. Kansagara, D., et al., *Risk prediction models for hospital readmission: a systematic review*. JAMA, 2011. 306(15): p. 1688-98.
35. Cook, D.A., *Methods to assess performance of models estimating risk of death in intensive care patients: a review*. Anaesth Intensive Care, 2006. 34(2): p. 164-75.
36. Lemeshow, S. and J.R. Le Gall, *Modeling the severity of illness of ICU patients. A systems update*. JAMA, 1994. 272(13): p. 1049-55.
37. Moons, K.G., et al., *Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker*. Heart, 2012. 98(9): p. 683-90.
38. *ACP Clinical Practice Guidelines*. [cited 2014 March]; Available from: http://www.acponline.org/clinical_information/guidelines/guidelines/.
39. Cook, N.R., *Use and misuse of the receiver operating characteristic curve in risk prediction*. Circulation, 2007. 115(7): p. 928-35.
40. *Number Crunching; Predictive analytics helps Presbyterian Healthcare analyze financial data, in Health Data Management*. 2010. p. 99.
41. Stevanovic, J., M.J. Postma, and P. Pechlivanoglou, *A systematic review on the application of cardiovascular risk prediction models in pharmacoeconomics, with a focus on primary prevention*. Eur J Prev Cardiol, 2012. 19(2 Suppl): p. 42-53.
42. Adelman, J.S., et al., *Understanding and preventing wrong-patient electronic orders: a randomized controlled trial*. J Am Med Inform Assoc, 2013. 20(2): p. 305-10.
43. Cruess, S.R. and R.L. Cruess, *The medical profession and self-regulation: a current challenge*. Virtual Mentor, 2005. 7(4).
44. *The Joint Commission*. [cited 2014 March]; Available from: http://www.jointcommission.org/.
45. *National Association for Healthcare Quality Professionals*. [cited 2014 March]; Available from: http://www.nahq.org/.
46. *The American Medical Informatics Association - About AMIA*. [cited 2014 March]; Available from: http://www.amia.org/about-amia.
47. Park, S.C., et al., *Validation of a scoring tool to predict drug-resistant pathogens in hospitalised pneumonia patients*. Int J Tuberc Lung Dis, 2013. 17(5): p. 704-9.
48. Yoon, Y.K., et al., *Clinical prediction rule for identifying patients with vancomycin-resistant enterococci (VRE) at the time of admission to the intensive care unit in a low VRE prevalence setting*. J Antimicrob Chemother, 2012. 67(12): p. 2963-9.
49. Breslow, M.J. and O. Badawi, *Severity scoring in the critically ill: part 1--interpretation and accuracy of outcome prediction scoring systems*. Chest, 2012. 141(1): p. 245-52.
50. Labarere, J., R. Bertrand, and M.J. Fine, *How to derive and validate clinical prediction models for use in intensive care medicine*. Intensive Care Med, 2014.
51. Hovenga, E.J., *Health workforce competencies needed for a digital world*. Stud Health Technol Inform, 2013. 193: p. 141-68.
52. Hart, M.D., *Informatics competency and development within the US nursing population workforce: a systematic literature review*. Comput Inform Nurs, 2008. 26(6): p. 320-9; quiz 330-1.
53. Hersh, W. and A. Wright, *What workforce is needed to implement the health information technology agenda? Analysis from the HIMSS analytics database*. AMIA Annu Symp Proc, 2008: p. 303-7.
54. Arnold, L.D., et al., *Statistical trends in the Journal of the American Medical Association and implications for training across the continuum of medical education*. PLoS One, 2013. 8(10): p. e77301.
55. Curran, C.R., *Faculty development initiatives for the integration of informatics competencies and point-of-care technologies in undergraduate nursing education*. Nurs Clin North Am, 2008. 43(4): p. 523-33, v.
56. Bates DW, S.S., Ohno-Machado L, Shah A, Escobar G, *Big Data In Health Care: Applying Analytics To Identify And Manage High Risk And High Cost Patients*. Accepted for Publication in Health Affairs, 2014.