At the turn of the last century, physicians were largely guided by lessons passed down in training, their own personal experience, and the experiences of their colleagues. Although this approach produced thoughtful clinicians, a key limitation remained—even the busiest, most experienced providers could see only so many patients, experience only a limited number of outcomes, and often struggled to ascertain the accuracy of diagnoses or the effectiveness of treatment. These challenges to delivery of safe and effective patient care were subsequently addressed by a growing focus on progressively larger and better designed cohort studies and randomized clinical trials and later by distilling these insights into clinical practice guidelines and appropriate use criteria to help summarize the rapidly evolving medical literature.

Article see p 368

Despite this exponential growth in well-conducted clinical research, a barrier in applying these studies into clinical care is that an individual patient may not obtain the average benefit observed in a clinical trial. It is well recognized that the heterogeneity of treatment effect across a population can be obscured by focusing only on the mean treatment effect in a population.1–3 Depending on a patient’s age, sex, comorbidities, and other characteristics, that patient may benefit greatly from the same treatment that poses a significant risk for another.4 Providers, after all, are concerned with delivering the safest and most effective treatment for a particular individual, rather than a population of patients. Accordingly, a growing focus has been placed on developing and implementing tools to identify which patients are likely to benefit from a particular treatment or strategy, those who may be harmed, and those for whom balanced risks and benefits exist that should alter a clinical decision. If a physician will not treat a low- or high-risk patient differently, then what value is there in using a model? With respect to development, the methods currently published are clearly wanting. Only a modest number of publications reported adequate quantification of the model’s capacity to discriminate the outcome of interest (just 63% reported a c-statistic) and even fewer reported a measure of calibration to help readers understand how the model’s predictions are compared with observed event rates (only 36% reported a Hosmer–Lemeshow statistic or a calibration plot)—the latter being critically important in prospectively using models to tailor treatment to risk. Moreover, only a minority of studies reported either an internal or external validation of the model. Another challenge, as the field strives to settle on the best models for clinical care, is that few models reported—just 3%—compared alternative prediction models.

Perhaps most distressing is the lack of clinical application of many of these models. Almost no studies have focused on the effective implementation of these models in clinical care—the principal goal of developing such tools. For example, although few cardiologists complete a day without using CHADS2 or CHADS2-Vasc models to predict stroke risk in patients with atrial fibrillation, there are few studies examining whether the use of these models improves the outcomes of patients with atrial fibrillation. One of the few examples of prospectively using a risk prediction model in clinical care, that we are aware of, is the use of the American College of Cardiology’s prediction model for peri–percutaneous coronary intervention bleeding, which was associated with a 44% reduction in the odds of bleeding.5–7 Implementing the best prediction models and demonstrating improvements in care are clearly a high priority for the profession and an important step toward precision medicine.

In this issue, Wessler et al report findings from a rigorous review of clinical prediction models in cardiovascular disease. Examining over 20 years of contemporary literature, they found ≈800 models focusing on conditions across the entire spectrum of cardiovascular disease. In fact, they found that the number of new clinical prediction models has doubled each decade. The growth of prediction models is encouraging, as the entire profession seeks to better understand the outcomes of their patients and how best to optimize these outcomes. Importantly, however, they highlight many important challenges, including the design, development, and testing of prediction models.

With respect to the design of clinical prediction models, it is critically important that the results of the model would alter a clinical decision. If a physician will not treat a low- or high-risk patient differently, then what value is there in using a model? With respect to development, the methods currently published are clearly wanting. Only a modest number of publications reported adequate quantification of the model’s capacity to discriminate the outcome of interest (just 63% reported a c-statistic) and even fewer reported a measure of calibration to help readers understand how the model’s predictions are compared with observed event rates (only 36% reported a Hosmer–Lemeshow statistic or a calibration plot)—the latter being critically important in prospectively using models to tailor treatment to risk. Moreover, only a minority of studies reported either an internal or external validation of the model. Another challenge, as the field strives to settle on the best models for clinical care, is that few models reported—just 3%—compared alternative prediction models.

Perhaps most distressing is the lack of clinical application of many of these models. Almost no studies have focused on the effective implementation of these models in clinical care—the principal goal of developing such tools. For example, although few cardiologists complete a day without using CHADS2 or CHADS2-Vasc models to predict stroke risk in patients with atrial fibrillation, there are few studies examining whether the use of these models improves the outcomes of patients with atrial fibrillation. One of the few examples of prospectively using a risk prediction model in clinical care, that we are aware of, is the use of the American College of Cardiology’s prediction model for peri–percutaneous coronary intervention bleeding, which was associated with a 44% reduction in the odds of bleeding.5–7 Implementing the best prediction models and demonstrating improvements in care are clearly a high priority for the profession and an important step toward precision medicine.

Although the article by Wessler et al in this issue, and the recent Prognosis Research Strategy (PROGRESS)5 and...
Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis (TRIPOD) statements underscore the need to elevate the methodological rigor of model development, there are many additional issues that need to be considered to better understand how prediction models to sort out heterogeneity of treatment benefit can be used to elevate the value of healthcare. First, models that have been shown to be effective deserve a more prominent focus in clinical practice guidelines. In their current form, guidelines focus largely on summarizing the mean treatment effect of trials and many times avoid endorsement of the most effective clinical prediction models to help personalize care. Although they often do include single-patient characteristics to stratify risk and recommend more effective therapies, the specificity of guidelines could be elevated by identifying valid models and recommending treatments for different thresholds of risk. This was recently adopted in the updated cholesterol guidelines, but its reach could be far greater and result in easier translation of the guidelines into clinical practice.

Once clinical prediction models have been shown to effectively stratify risk and use of the model has been shown to improve outcomes, understanding how to incentivize the model’s use is another challenge. For example, the Society of Thoracic Surgeons risk scores are easily accessible in the form of a web-based calculator. However, these risk scores were rarely used at the bedside until the introduction of transcatheter aortic valve replacement demanded their use in trial protocols and clinical guidelines. The other Society of Thoracic Surgeons models are still rarely used prospectively in routine clinical care at many institutions. Perhaps the most powerful incentive to better deploy risk models in practice is for payers to reimburse providers for using them—something that will also benefit the payers by allowing treatment for their patients to be preferentially provided to those who most benefit. Implementing risk models is complex, time consuming, and expensive, requiring iterative feedback to ensure optimal deployment and supporting the education of providers to know how best to interpret the results and share them with their patients. One potential avenue to support the use of risk models would be to link prospectively established risk, as advocated for by Diamond and Kaul. Such payments need not only be for physicians. For example, applying lower copays to patients who benefit substantially more from potent antiplatelet agents, such as ticagrelor or prasugrel, could support the use and adherence to such treatments. New and innovative strategies are needed, and research will be needed to identify additional techniques to enhance adoption of models at the point of care and support ongoing use.

Another largely unrealized opportunity for use of clinical prediction models is their use to support shared decision making between patients and physicians. Previous studies have found that a majority of patients prefer to play an active role in decision making at the time of acute myocardial infarction. Patients’ goals and preferences for care can vary widely, and the execution of risk models for individual patients could enable these patients to understand the benefits and risks of treatment for people like them. For example, a patient with a low SYNTAX score who is most concerned about avoiding stroke may prefer percutaneous coronary intervention to manage severe left main coronary disease, whereas another whose goal is to avoid returning to the hospital for additional procedures may strongly prefer bypass surgery. Using risk models to support these decisions represents the pinnacle of patient-centered care.

All of these potential applications of clinical prediction models demand continued creation and testing of these models. How will these efforts be incentivized and sustained? Although it is possible that some pharmaceutical and device manufacturers may fund these efforts so that they can create tools to support the use of their interventions at the time they publish the results of clinical trials, this has not yet been done. Professional societies, such as the American College of Cardiology and Society of Thoracic Surgeons, have built models, but these have been self-funded by-products of creating risk-standardized methods for comparing hospitals as part of a larger goal of improving quality. Governmental grants are unlikely to support such efforts, and the grant-writing effort often exceeds the workload in developing the models themselves. Although it is possible that models could be licensed and used to generate the resources needed to support their creation, this will prevent widespread access to potentially valuable models. Identifying funding mechanisms that support this important work is necessary for continued progress.

The future of clinical risk prediction models as a driver of more personalized care is promising. To realize this potential, it is critical that the outcomes research community continue to shift its focus from merely describing outcomes to creating models that can support the preferential use of treatments in patients who most benefit, while avoiding use in patients who do not benefit or might even be harmed. This will substantially increase the value of healthcare and is a priority for our field. Although the work of Wessler et al to assemble and critically evaluate existing models is an important first step, defining how best to disseminate, to implement and sustain the best models is a top priority.

Disclosures

Dr Spertus owns several patents on the ePRISM technology used to deliver clinical risk prediction models at the point of care and has an ownership interest in Health Outcomes Sciences, a company that distributes and supports the ePRISM software to hospitals. Dr Salisbury reports no conflicts.

References


Key Words: Editorials ◼ decision support techniques ◼ patient care ◼ physicians ◼ risk assessment
Realizing the Potential of Clinical Risk Prediction Models: Where Are We Now and What Needs to Change to Better Personalize Delivery of Care?

Adam C. Salisbury and John A. Spertus

Circ Cardiovasc Qual Outcomes. published online July 7, 2015;
Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2015 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circoutcomes.ahajournals.org/content/early/2015/07/07/CIRCOUTCOMES.115.002038.citation

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Quality and Outcomes can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Quality and Outcomes is online at:
http://circoutcomes.ahajournals.org//subscriptions/