Clinical epidemiology has sensitized primary care clinicians to the concept of risk factors for disease. Risk factors are patient characteristics that indicate increased likelihood of having or developing a disease or condition. These risk factors are commonly used by clinicians to stratify patients into groups that are at risk or at increased risk and those who are at low risk. For example, when patients have their blood pressure routinely assessed at a visit, the primary care clinician is assessing patients for the presence of hypertension, which is primarily measured as a risk factor for future cardiovascular disease events and mortality.\(^1\)

Unfortunately, many conditions have more than one risk factor that needs to be evaluated, thereby adding a degree of complexity to the assessment of risk. The evaluation of a single risk factor does not put that factor in the context of the patient’s other risk factors for disease, a job that routinely challenges clinicians. Consequently, there has been the development of “risk scores.”

A risk score takes many risk factors for an outcome and combines them into a single predictive measure. The human mind is unable to compute the relative influence of more than two or three risk factors at a time; beyond that, the mental calculus becomes too complex.\(^2\)

As an illustration, consider the following case:

Mrs C, a 62-year-old woman with hypertension that is controlled with medication (126/74 today), has a lipid profile measured. Her total cholesterol is 258 mg/dl, her low-density lipoprotein (LDL) cholesterol is 168 mg/dl, and her high-density lipoprotein (HDL) cholesterol is 68 mg/dl. She is a nonsmoker, and her family history is positive for myocardial infarction in her father.

This case presents a decision point that may be difficult for her clinician. How aggressive should the clinician be in suggesting cardiovascular risk reduction strategies? Should the clinician suggest a restrictive cholesterol-lowering diet? How much should the clinician stress weight reduction as a goal? Should this patient be started on a cholesterol-lowering medication to lower her risk for heart disease? In favor of therapeutic lifestyle changes and/or starting medication to lower her risk are her risk factors of diagnosed hypertension, elevated total and LDL cholesterol, and family history of heart disease. Factors that weigh against the cost, burden of monitoring, and potential side-effects of medication or lifestyle change are her elevated HDL cholesterol, her good blood pressure control, and her nonsmoking status.

Where is the balance between these factors, and if the balance weighs toward intervention, is lifestyle...
Risk assessment tools concurrently weigh the value of factors that increase and decrease the risk for an outcome and then succinctly summarize the combined risk for an individual as a risk score. A risk score weighs the value of each risk factor by assigning each of them a number of risk points. Risk points are then summed to yield a total score. Most risk scores are designed to score points for risk, although some contain both hazardous and protective factors, with protective factors scoring negative points to lower the composite risk score. For example, in a risk score predicting coronary atherosclerotic lesions in young people, being female has an associated risk of -1.0. Additionally, many risk scores then associate specific levels of the composite risk score with a probability for the disease outcome.

Many risk scores have as their objective to predict prospective risk of developing an outcome. For example, the Framingham Risk Score predicts an adult’s risk of having a heart attack in the next 10 years, based on historical and laboratory information. Another example of a prospective risk score is the Acute Physiology and Chronic Health Evaluation (APACHE) II risk score, which predicts risk of mortality for patients hospitalized in an intensive care unit. This score is often used in preoperative anesthesia consultations to evaluate perioperative risk. Another risk score predicts the 6-month mortality risk of nursing home residents with dementia and can be useful in making decisions about hospice care and in counseling about prognosis.

Alternatively, some risk scores assess a person’s current risk and are meant to stratify risk to identify those who may benefit from further screening for disease. These tools help clarify pretest probability, which can aid the clinician in making decisions about the next tier of screening and may also inform health policy. An example of this kind of risk score is the Danish Diabetes Risk Score, which was designed to identify those in the population most likely to have undiagnosed diabetes and who might be targeted for laboratory screening.

**Assessing the Quality of a Risk Score**

*Is It Useful for My Patient?*

In evaluating a risk score, the first consideration is, “How will this be used, and how will it be an improvement over what is currently available?” Thus, the type of risk assessment tool chosen depends on the context in which it will be used.

For a patient with known risk factors for coronary heart disease, such as in the case presented above, a risk score such as the Framingham risk score, with both historical and laboratory data is helpful and may aid the clinician in assisting the patient with secondary prevention efforts. In a different context, however, another risk score may be more appropriate, even for predicting risk for the same condition. For example, a risk score for coronary heart disease based only on self-reported historical data and indices that can be measured non-invasively, such as weight and blood pressure, might be used for initial risk assessment for an undifferentiated group of persons, perhaps in the clinician’s waiting room or at a health fair. The data contributing to this risk score are consistent with its purpose of population-based screening and may help identify people who might benefit from further evaluation.

Alternatively, some risk scores may be designed containing only factors that can be extracted from an electronic health record (EHR). These tools may be helpful for creating patient registries for research or quality improvement projects.

*Who Is Included?*

When assessing the value of a risk score as an aid to decision making for an individual patient, clinicians need to consider both the patient population from which the risk score was created and factors that might make a specific patient exceptional. The characteristics of the population from which the risk score is developed should be similar to the characteristics of the patient to whom the scoring system will be applied. Additionally, clinicians should keep in mind that risk scores are developed from data from large samples and predict the effect of risk factors on outcomes in heterogeneous populations. A risk score developed in a large sample or population, by definition, projects the “mean” outcome for that risk. Its behavior for individual patients may vary about the mean.

*What Risk Factors Were Assessed?*

Clinicians using risk assessment tools need to evaluate which risk factors are included in the score and, alternatively, evaluate if there are important risk factors that are missing. While this may seem self-evident, important risk factors for disease are often omitted if information about these risk factors is not available to the developers of the score. For example, in a risk score assessing risk for development of type 2 diabetes, known risk factors such as age, body mass index (BMI), and family history should be included. However, one would be hard-pressed to find a risk score that included other known diabetes risk factors such as polycystic ovarian syndrome (PCOS) or acanthosis nigricans (AN) as predictive factors. While these conditions carry known risk for diabetes, longitudinal data about people with PCOS and AN and development of diabetes is scarce.
The risk factors contained in the risk score need to be easily measured in the setting in which it will be applied. A risk score that uses a biomarker that is only available through a research laboratory is unlikely to be routinely used or worthwhile in clinical settings. When there are risks that can be measured in several different ways, the one that can be most easily measured is the most ideal, given comparable predictive abilities. For example, if BMI and triceps skin-fold thickness are equivalent predictors of diabetes, then BMI should be chosen for the risk score.

Is It Easy to Compute the Score?

Ideally, a risk scoring tool is conceptually straightforward and can be easily used and computed by clinicians in their daily practice. Resources available at the point of care, especially time, may be limited. For ease of use, many developers of risk scores round odds/hazard ratios from the multivariate model into whole-number risk points. Rounding to whole numbers is useful for mass screening using paper and pencil scoring. One could make the argument, however, that with increasing use of EHR, computers could easily calculate a risk score based on a non-integer values for risks. However, computers are still not readily accessible in all patient exam rooms or at risk-screening programs such as health fairs. Even if an EHR is in use, it must be programmed to perform the calculations needed.

With growing trends in the use of EHR and Internet resources, some newer risk scores do not convert the risks yielded in the multivariate risk prediction model to integers but rather use the standardized regression coefficients from the model and also keep transformed variables from the model. This strategy of creating scores based on the regression coefficients and transformed variables has the advantage of being more precise than rounding to the nearest whole number. The disadvantage to using regression coefficients and transformed variables is that, unless the risk score is already integrated into a computer program, it is difficult for the user to multiply six digit coefficients by the values for each risk factor (eg, natural log of HDL cholesterol) and then add it up.

Whichever method is used, once risk points are assigned, a total score can be computed. For most outcomes, it is clinically and conceptually useful to define one or more cut-off levels, above which the person is at increased risk for disease. While most risk scores will likely have a linear correlation between increasing score and increasing risk, receiver operating characteristic (ROC) curves can be used to define the value at which the balance between sensitivity and specificity is optimized. ROC curves plot a test’s true positive rate against its false positive rate (or its sensitivity against 1-specificity) and identify the cut-off score that gives the best sensitivity and specificity. Different cut-off values may be desired, however, depending on whether the user wants to increase the sensitivity of the measure or to optimize specificity. Stated differently, different cut-off values may be chosen depending on whether the desire is to minimize false negatives and thus identify nearly all who might have the condition (high sensitivity), or to minimize false positives and thus avoid needlessly worrying or testing those who are unlikely to have disease (high specificity).

Has the Risk Score Been Validated?

Evidence of validation of a risk score in a population other than the one in which it was created enhances the credibility of the score. One way to validate a risk score is to split the original population into two parts. In this split sample, one half (or another fraction) of the data set is randomly chosen for development of the risk score (the derivation or development set), and the risk score is then validated in the other half of the data (the validation set). One drawback to this approach is that if there are a limited number of people in the overall sample, splitting the sample in two may cause a loss of statistical power such that a good predictive model can no longer be created. An alternative to splitting the sample is “bootstrapping.” Bootstrapping is a statistical method in which multiple smaller samples are taken out of the whole sample, and the predictive model is tested in each of the smaller samples.

Beyond validating using a split sample or bootstrapping, the validity of the risk score is optimally demonstrated by validating in other populations. Validation in populations with different characteristics than the source population demonstrates that the risk score is generalizable and robust. It may be valuable to demonstrate that a risk score performs equally well in younger or older populations, one with a different racial or ethnic composition, or in primary care patients versus specialty patient populations. Several sources discuss more of the statistical concepts and techniques involved in creating and validating risk assessment tools.

Applying Risk Assessment Tools to Individual Patients

Returning to the case of Mrs C, her care provider may have difficulty balancing her risks against protective factors. A savvy clinician might use the Framingham risk score to evaluate Mrs C’s 10-year risk for heart disease. A clinician with Internet access might even use an automated risk score calculator, where getting a percent risk for heart disease is as easy as entering a few pieces of data from the above scenario and clicking “calculate.” For Mrs C, the Framingham 10-year risk for heart attack is estimated at 3%. The clinician with a computer in the exam room could go even further by turning the computer’s monitor toward the patient,
and involving the patient in the process of calculating risk. The clinician could show the effect of lowering cholesterol on the risk for heart disease, thereby using the Framingham risk calculator as both a decision aid and as a teaching tool for shared decision making.22

Clinicians can also use risk scores to stratify patients into low-, medium-, and high-risk groups to assess the pretest probability for disease. This can then influence the decision path that the physician takes in diagnostic testing and treatment. Consider the following case of a patient presenting with pharyngitis:

Patient P is a 16-year old boy who presents complaining of 2 days of sore throat, fever at home to 38.2°C, and a slight cough. He has no known ill contacts. His mother gave him acetaminophen to reduce fever 2 hours ago. Physical exam reveals a moderately ill-appearing boy without fever or rash. Eyes and ears are normal, but examination of the pharynx reveals tonsillar swelling with erythema and exudate. His neck has tender anterior cervical adenopathy.

This patient presentation prompts a decision of whether to test and to treat for streptococcal pharyngitis. His history of fever and physical exam findings suggest streptococcal pharyngitis, but his slight cough and lack of ill contacts weigh against the diagnosis of strep throat. The diagnosis is important because untreated streptococcal pharyngitis can not only result in a missed opportunity for the patient to get better in a few days with treatment but also in complications such as peritonsillar abscess and rheumatic fever. However, the approach of testing every patient presenting with pharyngitis symptoms using a rapid test for Group A streptococcal antigen is both costly and is likely to lead to overtreatment for false positive results and undertreatment for false negative results.23

How is the clinician to decide whether to perform a rapid streptococcal antigen test? Having an accurate assessment of the pretest probability of streptococcal pharyngitis can aid in clinical decision making by identifying a range in which testing is likely to be useful. Generally, the greatest knowledge from diagnostic testing, and the most benefit for patients, can be gained when pretest probability is near 50%. The Melsaac risk score for strep can help the clinician estimate pretest probability.5,24 Using this score, Patient P would score 1 point for fever, 1 point for pharyngeal exudate and swelling, and 1 point for tender anterior cervical adenopathy, for a total score of 3. A score of 3 corresponds to a pretest probability of 35%.5,23 This is close enough to 50% that it would probably be beneficial to test this patient. Conversely, when pretest probability is very high or very low, testing yields less information, and the probability of false negatives or positives, respectively, becomes an important issue.

Clinicians will also encounter risk scores in clinical research. For example, the Charlson Comorbidity Index, which predicts 1 year mortality, is frequently used in research as a summary measure to characterize the burden of comorbidities for individuals in a sample.14 Researchers use the Charlson Index in survival analyses to assist in adjusting for the potential confounding effect of other comorbid diseases.25 For example, when evaluating the effect of weight fluctuation on mortality, it is useful to consider the effect of other diseases such as diabetes and hypertension on mortality.26 In such cases, the Charlson Index can help the researcher isolate the effect of weight fluctuation.14,26

Further Considerations for Users of Risk Assessment Tools

While risk scores are attractive because of their ability to help clinicians and researchers synthesize the multiple risk factors that influence the occurrence of disease, they should not be used and “obeyed” with abandon and should not necessarily override clinical judgment. Risk assessment of individual patients is a complex integrative process that may not be adequately addressed by a risk scoring system. As with most tools, risk scores are most helpful when used properly.

A potential limitation of risk scores is changes in validity in different groups to which they might be applied. For example, the Framingham Risk Score can be used with men and women while the Reynolds Risk Score, which was created from a women-only cohort, can be applied only to women.4,15 If a risk score was developed in a population-based sample, does it maintain its validity for specific subgroups of that population? Conversely, and perhaps more commonly, if the risk score was developed in a specific group, for example a specialty clinic population, does it apply equally or acceptably well to a primary care patient population? The answer to this question may often be no, as the greater prevalence of selected conditions in specialty clinic populations will yield a higher positive predictive value for a test than the general population, where prevalence of disease is lower.27

Many risk scoring systems are available for clinical decision making. Clinicians need to evaluate how risk scoring systems might fit into their work flow and evaluate the trade-off between time spent in formal risk assessment versus the magnitude of the derived benefit. Barriers cited by primary care physicians to the use of risk scores included concerns that a single risk value might not accurately represent the complex patient, that important risk factors were not included in the risk scores, that prediction rules might result in overtreatment, and that application of prediction rules is time-consuming.28 The creation of Web-based calculators to assist in efficient calculation of risk scores and the integration of these calculators into EHRs can facilitate use.15,25
Conclusions
Risk scores are common and can assist in assessing a patient’s risk of having or developing disease. The most useful risk assessment tools will be those that address the most difficult and salient decision points for clinicians and patients, are easily incorporated into clinician workflow, and have been found to be robust in validation studies. The future of risk scores likely lies in their ability to be integrated into, and calculated by, the EHR for point-of-care reference and decision making.

Acknowledgments: This study was supported by grant 1D12HP00023 from the Health Resources and Services Administration.

We would like to thank David Mehr, MD, MS, of the University of Missouri-Columbia Department of Family and Community Medicine for his assistance with critical review of the manuscript.

Corresponding Author: Address correspondence to Dr Koopman, University of Missouri-Columbia, Department of Family and Community Medicine, MA306 Medical Sciences Building, DC032.00, Columbia, MO 65212. Fax: 573-884-6172. koopmanr@health.missouri.edu.

REFERENCES