TO THE EDITOR: It was with great interest that we read the recent article by Whiting and colleagues (1) introducing the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) tool, a revision of the tool from 2003 (2). It is important that such pivotal methodological work about diagnostic review research is made widely available through publication in Annals of Internal Medicine.

We would like to discuss some issues related to QUADAS-2, as some of them attracted our attention when reading the otherwise impressive article. First, in the original QUADAS tool (2), items 1 and 2 related to variability with the potential to affect clinical generalizability (external validity) of study results. The term “applicability” is commonly used as a synonym (3). But, in QUADAS-2, “applicability” refers to whether certain aspects of an individual study are matching or not matching the review question (1). During a systematic review, the review question (according to Participants, Interventions, Comparisons, Outcomes, and Setting [PICOS] [4]) and eligibility criteria are prespecified before systematically searching for relevant studies (5). Therefore, all studies reaching the stage of quality assessment in a systematic review should be applicable to the review question because studies not matching the review question and eligibility criteria had already been sorted out during the selection process. The use of applicability in QUADAS as the “degree to which the results of a study can be applied to patients in practice” (2) may thus be preferable.

Second, some aspects of QUADAS-2 seem to be more time-consuming (that is, drawing a flow chart for each study). What is the average time needed to evaluate a study using QUADAS-2 versus QUADAS?

Third, the examples of study assessment that would facilitate understanding of QUADAS-2 are not yet available from www.quadas.org. In addition, an explanation paper further describing use in detail would be very helpful for researchers conducting diagnostic accuracy meta-analyses. Is this planned?

Lastly, calculating the interrater reliability only for agreement on the domain level when piloting the tool may not be appropriate to detect variability on the level of signaling questions, which, however, are used for the important decision on whether high or low risk of bias is present in individual studies.

Sabine Schueler, MD
Georg M. Schuetz, MD
Marc Dewey, MD, PhD
Charite–Universitaetsmedizin Berlin
Berlin 10117, Germany

Potential Conflicts of Interest: Dr. Dewey: Consultancy: Guerbet; Grants/grants pending (money to institution): European Regional Development Fund, German Research Foundation/German Federal Ministry of Education and Research, GE Healthcare, Bracco, Guerbet, Toshiba Medical Systems; Payment for development of educational presentations including service on speakers bureaus (money to institutions): Toshiba Medical Systems, Bayer-Schering, Guerbet, Cardiac MR Academy Berlin; Travel/accommodations expenses covered or reimbursed (money to institutions): Toshiba Medical Systems, Bayer-Schering; Other (money to institution): Course Director for Cardiac CT Hands-on Courses at Charite in Berlin, institutional master research agreements with Siemens Medical Solutions, Philips Medical Systems, and Toshiba Medical Systems; Other: Book author for Springer titles Coronary CT Angiography (2009) and Cardiac CT (2011).

References

IN RESPONSE: We thank Drs. Schueler, Schuetz, and Dewey for their comments regarding our recent article and are grateful for the opportunity to comment on the important issues they raise.

It is not strictly correct that items 1 and 2 of the original QUADAS tool (1) relate to variability (external validity). Item 1 covers both bias in patient selection and variability of included patients, whereas item 2 relates solely to reporting quality. Distinguishing between risk of bias (internal validity) and applicability (external validity) was key to developing QUADAS-2. Assessment of applicability requires a clear definition of the patients, tests, and target condition of interest. Dr. Schueler and colleagues suggest that applicability should be assessed in relation to “patients in practice” rather than the review question, as all included studies should be applicable to the latter. If there is an exact match between the review question and inclusion criteria, then this is correct. However, in practice, inclusion criteria are often broader than the review question. We do not think it is feasible to assess applicability in relation to “patients in practice” because the test may be applied to multiple patient groups and there may be different test variants or definitions of the target condition.

Dr. Schueler and colleagues state that completing QUADAS-2 may be more time-consuming and asked about the average time to evaluate a study using QUADAS-2 compared with the original tool. A user survey (2) found that completing the original tool could take from less than 10 minutes to 1 to 2 hours. We do not have such data for QUADAS-2, but from experience, we have found that completing QUADAS-2 takes around 10 minutes in addition to data extraction, although this is likely to vary considerably on the basis of reviewer experience and review topic. There were no negative comments regarding the time taken to complete the assessment during piloting.

We hope to make worked examples of QUADAS-2 assessments and more detailed, domain-specific guidance available on www.quadas.org in the near future. An explanatory paper is not currently planned, although we are working on updating the quality assessment chapter in the Cochrane Handbook for Diagnostic Test Accuracy Reviews to include QUADAS-2; this will include more detailed guidance. We believe that agreement at the domain level is a key concern.
TO THE EDITOR: I read with great interest Lederle and colleagues’ systematic review (1) on venous thromboembolism (VTE) prophylaxis. The authors conclude that heparin prophylaxis resulted “in little or no net benefit.” This is in contrast to an earlier meta-analysis by Wein and colleagues (2) that used many of the same studies in medical patients and patients with stroke. I believe that Lederle and colleagues reached this conclusion erroneously.

First, the authors’ decision to eliminate asymptomatic deep venous thrombosis (DVT) from their analysis accounts for a substantial bias against the efficacy of VTE prophylaxis. The authors do not provide the total number of asymptomatic DVTs in the included studies. However, if the DVT outcomes from Wein and colleagues’ meta-analysis were used, the point estimate for the absolute effect per 1000 patients would be 62 fewer DVTs, which is 7- and 15-fold greater than the rates of all bleeding and major bleeding events, respectively, per 1000 persons in all medical patients and patients with stroke in Lederle and colleagues’ analysis.

Second, by eliminating the outcome of asymptomatic DVT, the authors underestimate the effect of prophylaxis on symptomatic VTE. Because trial patients who developed asymptomatic DVT probably would have had treatment with systemic anticoagulation, it is reasonable to expect that this early detection may have reduced the rate of symptomatic events. This is strongly suggested by the fact that the rate of symptomatic DVT in Lederle and colleagues’ analysis is actually lower than the rate of symptomatic pulmonary embolism (PE), a finding that differs substantially from that seen in other trials targeting only symptomatic end points (3, 4).

Third, the authors emphasize some of their adverse effect data while simultaneously deemphasizing VTE outcome data. For medical patients, the point estimates for the absolute effect per 1000 persons suggest 10 fewer deaths, PEs, or symptomatic DVTs with only 1 major bleeding event. I think that most patients and providers would accept this tradeoff. Furthermore, the increase in all bleeding events with heparin is overwhelmingly minor, as shown by a major bleeding rate that is one tenth the rate of all bleeding events. If the authors are prepared to discount asymptomatic DVT, they should be willing to discount minor bleeding events as well.

I sincerely hope that this misleading systematic review does not impair the many efforts being undertaken around the world to improve the use of VTE prophylaxis in hospitalized patients.

Todd E.H. Hecht, MD
Raymond and Ruth Perelman School of Medicine, University of Pennsylvania
Philadelphia, PA 19104

References
scanning or pulmonary angiography, as quoted by Wiener and colleagues (6). Weiner and colleagues’ article also generates questions about the need to treat asymptomatic PEs, casting doubt even on the “strong” second recommendation—that pharmacologic prophylaxis with heparin or related drugs “should” be used in at-risk medical patients and those with stroke if bleeding risk is low.

These articles and guidelines also generate interest in possible new ideas. Do we need to take a different approach in our aggressiveness in dealing with at least some of the categories of DVTs, such as those that occur below the knee (and perhaps upper-extremity DVTs and central line–related DVTs, too)? In fact, many physicians have reservations about treating below-knee DVTs. As we are all aware, these DVTs are, at present, recommended to be treated like above-knee DVTs: with full anticoagulation.

So, we need to move forward from questioning the sweeping recommendation for VTE prophylaxis—which these articles support—to questioning the recommendations for therapeutic anticoagulation of low-risk DVTs (below the knee) and PEs (those that are incidental and asymptomatic).

Raghavendra R. Allareddy, MD, MBBS
Cleveland Clinic
Cleveland, OH 44195

Potential Conflicts of Interest: None disclosed.

References

IN RESPONSE: Dr. Hecht objects to our decision not to include asymptomatic DVTs in our systematic review. As noted in our introduction, asymptomatic DVT is a surrogate outcome that is more common than clinically evident VTE, and its value has been repeatedly questioned. We stand by our decision but agree that screening for asymptomatic DVT, which occurred in some trials in our review, may have reduced the number of symptomatic DVTs.

Treating asymptomatic DVT is analogous to increasing the crossover rate and, thus, could reduce the effect of prophylaxis on all measures of benefit and harm, including PE and bleeding events. Although we believe the practice of treating asymptomatic DVT in these trials to be misguided, we do not consider it a reason to include a surrogate outcome in our review. To assess the effect of treating asymptomatic DVTs on our findings, we compared the effect of heparin prophylaxis on asymptomatic DVT in studies that screened for asymptomatic DVT versus those that did not (1–6).

The effect did not differ substantially or significantly ($P = 0.83$) in medical patients (Figure), which suggests that this practice did not affect our findings much. The data are insufficient to conduct a similar analysis in patients with stroke. The Figure includes symptomatic DVTs from Mahe and colleagues (3) that were omitted from the original review. With their inclusion, the odds ratio in the original table under “Heparin vs. no heparin” for symptomatic DVT in medical patients changes slightly to 0.83 (95% CI, 0.55 to 1.26), with the same absolute effect per 1000 patients treated of $−2$ (CI, $−5$ to 3).

Dr. Hecht also criticizes our emphasis on outcomes in medical patients and bases this criticism on differences in mortality and symptomatic DVTs that were not significant and a difference in PE for which evidence of publication bias was detected. Weighing these uncertainties against the significant increase in bleeding events, we stand by our characterization of heparin prophylaxis as demonstrating little or no net benefit. Furthermore, a recently published randomized trial of 8307 medical patients reported no trend toward reduced mortality or morbidity from heparin prophylaxis (7), which strongly supports our conclusion.

We agree with Dr. Hecht’s suggestion that the most appropriate comparison of disease rates is between symptomatic VTE and major bleeding events. To this end, we twice cited Carrier and colleagues’ article (8), which provides quantitative support for that comparison.

We cannot dispute Dr. Allareddy’s point that recommending prophylaxis “if bleeding risk is low” is “questionable,” because, as noted in our discussion, we did not find trial data that identified patient subgroups more likely to benefit from prophylaxis, with the possible exception of persons older than 75 years.

Frank A. Lederle, MD
Roderick MacDonald, MS
Timothy J. Wilt, MD, MPH
Minneapolis Veterans Affairs Medical Center
Minneapolis, MN 55417

Potential Conflicts of Interest: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M10-1375.

References
Letters

Figure. Symptomatic DVT.

<table>
<thead>
<tr>
<th>Heparin</th>
<th>No Heparin</th>
<th>Weight, %</th>
<th>Peto Fixed Odds Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome and Study, Year (Reference)</td>
<td>Events, n</td>
<td>Total, n</td>
<td>Events, n</td>
</tr>
<tr>
<td>Did not screen for asymptomatic DVT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lederle et al, 2006 (1)</td>
<td>5</td>
<td>140</td>
<td>8</td>
</tr>
<tr>
<td>Weber et al, 2008 (2)</td>
<td>1</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Mahé et al, 2005 (3)</td>
<td>19</td>
<td>1230</td>
<td>21</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>1380</td>
<td></td>
<td>1394</td>
</tr>
<tr>
<td>Total events</td>
<td>25</td>
<td></td>
<td>29</td>
</tr>
<tr>
<td>Heterogeneity: chi-square = 1.53 (P = 0.46); I² = 0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 0.52 (P = 0.60)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did screen for asymptomatic DVT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohen et al, 2006 (4)</td>
<td>0</td>
<td>429</td>
<td>0</td>
</tr>
<tr>
<td>Samama et al, 1999 (5)</td>
<td>9</td>
<td>731</td>
<td>4</td>
</tr>
<tr>
<td>Leizorovicz et al, 2004 (6)</td>
<td>10</td>
<td>1856</td>
<td>15</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>3016</td>
<td></td>
<td>2641</td>
</tr>
<tr>
<td>Total events</td>
<td>19</td>
<td></td>
<td>19</td>
</tr>
<tr>
<td>Heterogeneity: chi-square = 0.57 (P = 0.45); I² = 0%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 0.71 (P = 0.48)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Test for subgroup differences: chi-square = 0.05 (P = 0.83); I² = 0%

DVT = deep venous thrombosis.

Preventing Unintended Consequences of Quality Measurement for Venous Thromboembolism Prophylaxis

TO THE EDITOR: Venous thromboembolism (VTE) remains a common complication of hospital care (1, 2). Multiple medical record audits of high-risk medical and surgical patients have consistently shown underuse of VTE prophylaxis (2). Most patients who are sick enough to require hospitalization have risk factors for VTE. Pulmonary embolism is a common and potentially preventable cause of death from hospital care (1, 2).

In their editorial, Baker and Qaseem (3) criticize The Joint Commission’s performance measure on VTE prophylaxis because it purportedly encourages its use for hospitalized patients unless there is documentation in the medical record that the patient is not at risk for VTE. We believe that this measure construct, using a permissible exclusion, remains appropriate. In 2006 and 2007, the National Quality Forum (NQF) convened a steering committee and a technical advisory panel to evaluate existing measures and develop new ones to address prevention and treatment of VTE (4). These panels included representation from multiple stakeholders, including 2 authors of the American College of Chest Physicians (ACCP) clinical practice guideline. The measures were tested by The Joint Commission in multiple hospitals and were posted for comment during the consensus development process at the NQF.

The decision to promote routine VTE prophylaxis or require documentation that no prophylaxis was needed was based on careful consideration of the ACCP guideline recommendations. The ACCP did not advocate patient-specific assessment for VTE risk due to the lack of appropriately derived and validated risk assessment tools, and stated that “there is sufficient evidence to recommend routine thromboprophylaxis for most hospitalized patient groups” (2). It also noted that “we are not able to confidently identify the small population of patients in the various groups who do not require thromboprophylaxis” (2).

Although the ACCP highlighted the lack of clinically validated risk assessment models to guide thromboprophylaxis, a careful review of the performance measure specifications for VTE-1 shows that the measure construct allows documentation of the lack of need for VTE prophylaxis based on completed hospital risk assessment forms (5). These specifications also give complete discretion to clinicians at the bedside to decide if prophylaxis is indicated for their patient and to
Potential Conflicts of Interest: Dr. Bratzler: Consultancy (money to institution): Johnson & Johnson Healthcare, Medline Industries. Dr. Loeb: Employment: The Joint Commission; Grants/grants pending (money to institution): The Joint Commission; Payment for lectures including service on speakers bureaus (money to institution): The Joint Commission. Dr. Kruenoski: Employment: The Joint Commission; Grants/grants pending (money to institution): The Joint Commission; Payment for lectures including service on speakers bureaus (money to institution): The Joint Commission.

References

TO THE EDITOR: In recent years, pay-for-performance programs and public reporting of provider performance have become integral components of strategies to improve the quality of health care in many countries. However, as highlighted by Baker and Qaseem (1), quality measurement can also have unintended consequences. Among the measures that these authors advocate to reduce such consequences is allowing physicians to report exceptions for not following a recommendation. Although exception reporting should be an essential component of quality measurement to ensure that patients who will not benefit from an intervention (such as those who are terminally ill) are not treated inappropriately, it is also important to monitor exception reports at both patient and provider levels.

In the United Kingdom, the Quality and Outcomes Framework, a major pay-for-performance system introduced into primary care in 2004, permits primary care physicians to report exceptions from quality targets in the framework (2). One potential disadvantage of allowing exception reporting is that physicians might exclude patients in whom they have difficulty reaching targets to increase their quality achievement scores. Excluded patients may then receive suboptimal care, resulting in pay-for-performance programs and quality measurement worsening existing disparities in health care if already disadvantaged groups of patients are more likely to be excluded from measurement of achieved quality targets.

Analysis of data on rates of exception reporting at the provider level in the Quality & Outcomes Framework shows a wide variation among primary care practices, particularly for indicators more closely linked to clinical outcomes, such as hemoglobin A1c control in people with diabetes (3). At the patient level, when we examined rates of exception reporting among people with diabetes from 23 primary care practices in socioeconomically and ethically diverse part of London, we found that patients who were exception-reported by their physicians were significantly less likely to achieve treatment targets for hemoglobin A1c, blood pressure and cholesterol control (4). There was considerable variation in exception reporting by patient characteristics, such as age, ethnicity, comorbidity, and duration of illness. The highest rates of exception reporting occurred among already disadvantaged groups at greatest risk for diabetes complications and poorer health outcomes, such as older patients, patients from ethnic minority groups, and patients with longer durations of diabetes and increased levels of comorbidity.

This experience from the United Kingdom highlights the importance of monitoring exclusions from pay-for-performance and public reporting programs. Failure to adequately monitor exception reporting and the reasons why physicians and providers exclude patients from public reporting may lead to exclusion of patients who are most at risk for adverse health outcomes, thus preventing them from receiving evidence-based care (5). This in turn may exacerbate current health inequalities and worsen health outcomes in vulnerable patients.

Azeem Majeed, MD
Imperial College London
London W6 8RP, United Kingdom

Potential Conflicts of Interest: Dr. Majeed is a general practitioner in the primary care practice of Dr Curran & Partners in London, United Kingdom. The Department of Primary Care & Public Health at Imperial College London has received funding from the NHS for research on pay-for-performance.

References
blood pressure or hemoglobin A1c. Reaching recommended targets should help discourage attempts at gaming.

outliers and to assess the validity of the exceptions reported, which should be recorded in clinical practice guidelines. This will make it easy and practical to audit exceptions. However, as we all agree, there are no established prediction rules to identify patients who do not benefit from VTE prophylaxis, and there are no guidelines from national experts on who should be considered low-risk. It therefore seems imprudent to rely on every hospital to develop its own “risk assessment form” to identify patients for whom VTE prophylaxis is not indicated. A better alternative would be to use a more restricted measure that applies only to patients for whom there is general agreement that VTE prophylaxis is indicated. For example, the ACCP recommends VTE prophylaxis for “acutely ill medical patients admitted to hospital with congestive heart failure or severe respiratory disease, or who are confined to bed and have one or more additional risk factors, including active cancer, previous VTE, sepsis, acute neurologic disease, or inflammatory bowel disease”(1). Taking the position that all patients should receive prophylaxis unless an undefined exception is present is not consistent with the evidence found by the American College of Physicians’ clinical practice guideline (2), and doing so may have the unintended consequence of encouraging the use of VTE prophylaxis for some low-risk patients for whom the risk–benefit ratio is unfavorable.

We agree with Dr. Majeed that although exception reporting is essential to discourage unnecessary or inappropriate treatment or testing, the rate of exception reporting should be routinely monitored to detect high outliers. As electronic health records become more widespread, exceptions should be routinely recorded in clinical decision-support tools; this will make it easy and practical to audit outliers and to assess the validity of the exceptions reported, which should help discourage attempts at gaming. Nevertheless, exception reporting will always require a careful balance, particularly for measures of intermediate outcomes, such as blood pressure or hemoglobin A1c. Reaching recommended targets for these measures requires both proper medical care and an informed, activated patient who is not hindered by financial or other obstacles. Simply allowing clinicians to document exceptions for anyone who cannot easily meet a recommended goal could result in worse care for the most vulnerable patients. However, not allowing exception reporting could encourage clinicians to push patients out of their practice if they are unable or unwilling to follow recommended regimens.

Amir Qaseem, MD, PhD, MHA
American College of Physicians
Philadelphia, PA 19106

Potential Conflicts of Interest: Dr. Qaseem was an author of the American College of Physicians’ guideline on VTE prophylaxis (2). Forms can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M11-2065.

References

A Lesson From the Slaughterhouse

TO THE EDITOR: I read the observation from Pusl and colleagues (1) regarding exogenous thyrotoxicosis with interest. The great medical journalist Berton Roueche described a similar phenomenon in his long-running column in The New Yorker, “Annals of Medicine” (2). In 1985, investigators from the Centers for Disease Control and Prevention traced an epidemic of thyrotoxicosis in a South Dakota meatpacking plant. Neck trimmings from the plant were being sold as 90% lean ground beef. The link was proven by, after approval by the Human Subjects Committee, having 4 volunteers eat cooked samples of the meat. Blood samples were obtained before and after consumption. Postconsumption samples showed the presence of thyroid hormone. This epidemic led the U.S. Department of Agriculture to prohibit meat processors from trimming near the gullet.

David Shore, the cocreator of the television series “House,” has said that the lead character of Gregory House was inspired in part by Roueche’s writings (3).

Victor S. Sloan, MD
Robert Wood Johnson Medical School
Flemington, NJ 08822

Potential Conflicts of Interest: None disclosed.

References

Correction

Correction: Photographer Name

The name of the photographer of the cover photograph on the 20 December 2011 issue was incorrect. The photographer’s name is Jitesh Kar, MD.