TO THE EDITOR: In their recent article on preoperative anticoagulant activity of low-molecular-weight heparin (LMWH) after a standardized bridging regimen, O’Donnell and colleagues (1) indicated the following: warfarin was withheld 4 to 5 days before the planned procedure; enoxaparin, 1 mg/kg of body weight twice daily, was given to patients 3 days before surgery; and the final dose of enoxaparin was given the night before surgery. They conclude that because the mean anti–factor Xa level before surgery was 0.6 U/mL and anti-Xa levels were 0.5 U/mL or greater in 68% of the patients, preoperative bridging therapy will result in a high residual anti-Xa level if the last dose is given the evening before surgery. They recommend that if a twice-daily regimen of LMWH is used, the final dose should be given in the morning of the day before surgery. Because anticoagulant-related postoperative bleeding rates did not increase in their patients, I do not think that such a recommendation can be made merely because the residual anti-Xa levels before the procedure were high.

The therapeutic range of anti-Xa for monitoring LMWH has not been established. The product insert for enoxaparin states that for a dosage of 100 U/kg (1 mg/kg) twice daily, the expected anti-Xa level at 3 to 4 hours after administration is less than 1.15 U/mL. Hirsh and Raschke (2) state that “a conservative therapeutic range for peak effect with twice-daily administration of enoxaparin . . . is 0.6 to 1.0 IU/mL.” Anti-Xa heparin levels measured by different laboratories vary greatly (3). In the College of American Pathologists survey (3), 87 laboratories used the chromogenic assay and LMWH as a standard for assay of the same sample. The mean anti-Xa level by these laboratories was 1.427 U/mL (SD, 0.449) (range, 0.34 to 2.60 U/mL). O’Donnell and colleagues report that the mean middosing anti-Xa level was 1.3 U/mL. Using twice-daily enoxaparin (1 mg/kg) and chromogenic assay for anti-Xa heparin level (the same as O’Donnell and colleagues) and aiming for an anti-Xa level of 0.5 to 1.0 U/mL, Ferreira and coworkers (4) reported that the mean middosing anti-Xa heparin level was 0.58 U/mL for 67 patients.

If O’Donnell and colleagues believe that the mean residual heparin level of 0.6 U/mL before surgery is too high, then they must also conclude that the enoxaparin dose of 1 mg/kg twice daily was too high. The mean anti-Xa heparin level of 1.3 U/mL at the middosing point that they found is higher than recommended therapeutic range for enoxaparin.

The method that I have used successfully for the past 10 years for preoperative management of anticoagulation seems to be simpler and to require less LMWH after temporary withdrawal of warfarin therapy. Patients are instructed to continue with warfarin therapy until the day before surgery. On the day before surgery, the patient comes to the clinic in the morning for an international normalized ratio check. Depending on the international normalized ratio, the patient is given 1 to 2 mg of vitamin K (using a parenteral preparation) orally, while the patient is in the clinic. This dose of vitamin K is adequate to bring the international normalized ratio to less than 1.5, which is safe for general surgery. The patient is then told not to take any warfarin until after surgery. After surgery, the patient starts taking warfarin as soon as he or she can take oral feeding and is given dalteparin (200 U/kg daily subcutaneously) 8 hours after surgery if he or she has no active bleeding. This method not only requires shorter duration of LMWH therapy but also reduces the chance of an unnecessary interruption of warfarin preoperatively in case, for any reason, a patient’s surgery is cancelled.

A.M. Shojania, MD
St. Boniface General Hospital
Winnipeg, Manitoba R2H 2A6, Canada

Potential Financial Conflicts of Interest: None disclosed.

References

IN RESPONSE: We appreciate the opportunity to respond to issues raised by Dr. Shojania. We agree that the therapeutic anti-Xa heparin level range for LMWH has not been rigorously evaluated; however, the high residual anti-Xa heparin activity observed in our study is a cause for concern because it may increase bleeding during surgery. We believe that when physicians stop warfarin therapy before surgery, they do not intend to substitute another anticoagulant during the surgical procedure.

The difference in mean anti-Xa heparin levels between our study and the study by Ferreira and colleagues (1) is not surprising because their patient population was younger and, in response to anti-Xa measurements, the enoxaparin dose was adjusted to target anti-Xa heparin levels of 0.5 to 1.0 U/mL. Although the preoperative management outlined by Dr. Shojania sounds reasonable, we are aware of no published cohort studies or clinical trials evaluating the safety of this approach.

Martin J. O’Donnell, MB
Clive Kearon, MB, PhD
Alexander G. Turpie, MB
McMaster University
Hamilton, Ontario L8V 1C3, Canada

Potential Financial Conflicts of Interest: None disclosed.

Reference
Elevated Creatinine Levels and Quality of Care in Heart Failure

TO THE EDITOR: I read the recent article by Baker and colleagues (1) with interest. I was concerned that end-stage renal disease and chronic renal insufficiency were listed as valid exclusions for why individuals are not receiving an angiotensin-converting enzyme (ACE) inhibitor or angiotensin-receptor blocker (ARB) in heart failure. This is not consistent with current guidelines, and the Agency for Healthcare Research and Quality recommendations state that an elevated creatinine level is not an exclusion (2). Subgroup analyses of studies do not suggest a lower benefit of ACE inhibitors or ARBs (3) and as individuals with an elevated creatinine level are at higher risk for events, the argument can be made that the potential benefit of ACE inhibitors is greater for those with elevated creatinine levels.

Linda Fried, MD, MPH
Veterans Affairs Pittsburgh Healthcare System
Pittsburgh, PA 15240

Potential Financial Conflicts of Interest: None disclosed.

References

IN RESPONSE: We agree that severe renal insufficiency is not a contraindication to using an ACE inhibitor or ARB for patients with heart failure. However, some of these patients will have a severe decline in renal function after starting ACE inhibitor or ARB therapy, requiring discontinuation of treatment. Note that if a patient with severe renal insufficiency tolerates ACE inhibitor or ARB treatment, the patient is represented in both the numerator and the denominator. A physician can “get credit” for treating such a patient successfully but does not “fail” the measure if a patient does not tolerate the treatment.

How should performance measures account for comorbid conditions, such as renal insufficiency? Performance measures frequently allow patients who are not receiving guideline-concordant care to be excluded when certain comorbid conditions are present, even if these conditions are not absolute contraindications. Thus, in the case cited above, any patient not prescribed an ACE inhibitor or ARB with certain renal failure diagnosis codes would be excluded from the denominator of the quality measure, although only a subgroup might have sufficient medical justification for withholding treatment. Similarly, patients with heart failure who are not given a β-blocker would be excluded if asthma or chronic obstructive pulmonary disease were present even though only few patients may have bronchospasm severe enough to outweigh the benefits of treatment. This approach is also problematic because listing a condition as an exclusion may be taken to imply that physicians should not even try to initiate therapy.

Quality measures that rely on these kinds of simplifications are helpful up to a point. When care is only fair, there is room for improvement even among the simple cases. But as care improves, the utility of this kind of measure becomes more limited. To improve quality to very high levels, quality measurement needs to be able to determine whether good care is given to more complex patients. In recognition of the problems of using comorbid conditions as absolute exclusion criteria, the Physician Consortium for Performance Improvement currently recommends using a general “medical reason” exclusion, rather than a comprehensive list of relative contraindications or conditions requiring cautious prescribing. When reporting exclusions, physicians are encouraged to explicitly record in the medical record the reason why a therapy is not given. Electronic health record vendors should create standard methods for physicians to easily record justifications for their clinical decisions to improve the accuracy of quality measurement and minimize time burdens.

David W. Baker, MD, MPH
Stephen D. Persell, MD, MPH
Karen S. Kmetik, PhD
Feinberg School of Medicine, Northwestern University
Chicago, IL 60611

Potential Financial Conflicts of Interest: None disclosed.

New Tests for the Diagnosis of Latent Tuberculosis Infection

TO THE EDITOR: We read Menzies and colleagues’ meta-analysis of new tests for latent tuberculosis infection (1) with interest and generally concurred with the authors’ inferences for research and practice. As the target condition in the review lacks a gold standard, we were dissatisfied with the use of sensitivity and specificity—measures of test accuracy that cannot be computed without verification of disease status—in data synthesis. It is important for researchers to establish the evidence base for incorporation of new tests in clinical practice concerning latent tuberculosis infection by using sound methods for evaluating tests without a gold standard (2), and we would like to highlight some key issues important for such research.

Half a century ago, sensitivity and specificity were introduced to express diagnostic accuracy (3, 4) in studies where findings of medical tests were compared between patients known to have the disease of interest and those without disease by using a 2× 2 table that cross-classified index test and gold standard results. Nowadays, these measures of test accuracy are so ubiquitous that researchers often feel compelled to estimate them even when the disease status of study participants cannot be easily verified. In this situation, estimates of diagnostic accuracy are biased because standards for verifying the disease are lacking (2). No readily applicable gold standard is available for latent tuberculosis infection; therefore, ascertaining sensitivity and specificity directly is impossible. Menzies and colleagues (1) used surrogate gold standards, taking active tuberculosis cases to estimate sensitivity and low-risk persons to estimate specificity. This approach has many problems, some of which they acknowledge. Because data for the 2 accuracy measures are derived from different sources, the relationship between sensitivity and specificity cannot be

Karen S. Kmetik, PhD
Feinberg School of Medicine, Northwestern University
Chicago, IL 60611

Potential Financial Conflicts of Interest: None disclosed.

New Tests for the Diagnosis of Latent Tuberculosis Infection

TO THE EDITOR: We read Menzies and colleagues’ meta-analysis of new tests for latent tuberculosis infection (1) with interest and generally concurred with the authors’ inferences for research and practice. As the target condition in the review lacks a gold standard, we were dissatisfied with the use of sensitivity and specificity—measures of test accuracy that cannot be computed without verification of disease status—in data synthesis. It is important for researchers to establish the evidence base for incorporation of new tests in clinical practice concerning latent tuberculosis infection by using sound methods for evaluating tests without a gold standard (2), and we would like to highlight some key issues important for such research.

Half a century ago, sensitivity and specificity were introduced to express diagnostic accuracy (3, 4) in studies where findings of medical tests were compared between patients known to have the disease of interest and those without disease by using a 2× 2 table that cross-classified index test and gold standard results. Nowadays, these measures of test accuracy are so ubiquitous that researchers often feel compelled to estimate them even when the disease status of study participants cannot be easily verified. In this situation, estimates of diagnostic accuracy are biased because standards for verifying the disease are lacking (2). No readily applicable gold standard is available for latent tuberculosis infection; therefore, ascertaining sensitivity and specificity directly is impossible. Menzies and colleagues (1) used surrogate gold standards, taking active tuberculosis cases to estimate sensitivity and low-risk persons to estimate specificity. This approach has many problems, some of which they acknowledge. Because data for the 2 accuracy measures are derived from different sources, the relationship between sensitivity and specificity cannot be

Karen S. Kmetik, PhD
Feinberg School of Medicine, Northwestern University
Chicago, IL 60611

Potential Financial Conflicts of Interest: None disclosed.
studied, which is a critical step in evaluating threshold effect and making decisions about pooling data in meta-analyses (5). By definition, latent tuberculosis is a condition in otherwise healthy persons; therefore, accuracy estimates derived from these data are likely to misrepresent the true performance of the index tests in latent tuberculosis. Mixing different conditions into the data set introduces spectrum bias (6). Thus, neither sensitivity in active tuberculosis nor specificity in apparently healthy, low-risk persons without verification of absence of disease can be generalized for application in practice concerning latent tuberculosis infection.

An approach to index test validation that goes beyond the diagnostic accuracy paradigm (2) is needed for an alternative evaluation process in the absence of a gold standard. A recently published health technology assessment report provides such an approach to evaluate tests for latent tuberculosis infection (7). Because there is no gold standard for latent tuberculosis infection with which to assess the comparative accuracy of the tuberculin skin test and interferon-γ assays, the validation process should explore meaningful relationships between index test results and other test results and clinical characteristics. For example, the risk for latent tuberculosis infection is greatest among tuberculosis contacts who share a room with the index case for the longest time, and a validation study can evaluate whether results of tests for latent tuberculosis infection correlate with the level of exposure in an outbreak investigation that ascertains exposure to tuberculosis, performs various index tests of interest in eligible persons, and compares test results with exposure status. The role of new tests versus tuberculin skin tests may be evaluated further through studies that compare the likelihood of false results in bacille Calmette–Guérin vaccination and HIV infection. Other validation may come from observational studies of whether new test results are predictive of development of active tuberculosis in the future. Menzies and colleagues (1) touch on these aspects, but their meta-analysis focuses on sensitivity and specificity.

The results of validation studies can be systematically reviewed, and meaningful inferences can be drawn about the value of tests without using sensitivity and specificity. The apparent attraction of these 2 simple statistics makes it hard for us to give them up, but when a gold standard does not exist, we have no choice but to abandon them. This is a key recommendation that investigators in diagnostic research of latent tuberculosis infection must consider.

IN RESPONSE: We agree with Drs. Kunst and Khan that the lack of a proper gold standard is a fundamental problem of all cross-sectional studies of diagnostic tests for latent tuberculosis infection. We state this problem explicitly several times in our paper. We believe that longitudinal studies following cohorts of persons with positive or negative test results will be most valuable, because the later development of active tuberculosis is the only certain indicator of the presence of latent tuberculosis infection. Because treatment reduces incidence of disease, ideally, such cohorts of individuals would be untreated, which poses serious ethical issues. However, as we have pointed out elsewhere (1), individuals with discordant test results could be left untreated, as there is equipoise regarding their management, and prognosis of discordant results is the most critical issue for understanding the predictive value of interferon-γ release assays. Several prospective studies are currently being conducted in different settings (2, 3); we await these results with interest.

With regard to gradients of exposure, we reviewed all available studies (see our Table 3). However, the measurement and categorization of exposure, and disease in the source cases, were too heterogeneous to allow their integration for a proper meta-analysis. Systematic reviews and meta-analysis must take advantage of published literature to be informative. Hence, we assessed sensitivity and specificity, but pointed out clearly in the paper that these were surrogate measures with significant limitations. If we had only included published studies with the correct gold standard for latent tuberculosis infection (as above), there would have been no papers on which to base our estimates. A considerable amount of published literature is currently available, which has compared 2 or even all 3 currently available tests for latent tuberculosis infection. Both interferon-γ release assays reviewed are currently licensed in many countries and are actively marketed in North America and Europe. Therefore, their relative performance in different patient populations and clinical situations is of considerable interest. We feel strongly that ignoring this large body of information, because of certain limitations, would do a disservice to public health and clinical practitioners who are faced with making choices and managing patients now.

References

3. Ledley RS, Lusted LB. Reasoning foundations of medical diagnosis; symbolic logic, probability, and value theory aid our understanding of how physicians reason. Science. 1959;130:9-21. [PMID: 13668531]

Potential Financial Conflicts of Interest: None disclosed.

Dick Menzies, MD, MSc
Madhukar Pai, MD, PhD
Montreal Chest Institute, McGill University
Montréal, Québec H2X 2P4, Canada

Potential Financial Conflicts of Interest: None disclosed.
TO THE EDITOR: The authors of the review papers on aspirin (1) and nonsteroidal anti-inflammatory drugs (NSAIDs) (2) for the primary prevention of colorectal cancer ignore important limitations of observational studies; raise concerns that are unlikely to be valid (1); and fail to include the analysis on NSAIDs from the Physicians’ Health Study (PHS) (3), which was published in the interval covered. The authors correctly discuss shortcomings of randomized trials showing no protection in men and women—that is, low dose and short duration (1). They fail, however, to address shortcomings of observational studies on regular long-term drug use—that is, unmeasured confounding—as an alternative explanation for the reduced risk for colorectal cancer observed in most of these studies.

Specifically, regular long-term drug use is associated with healthy characteristics that are difficult to measure. These healthy characteristics of long-term, adherent users may be associated with reduced risks that are independent of drug effects, as evinced by reduced risks for many adverse outcomes in persons who adhere to placebo. This can lead to paradoxical relations in observational studies (4). To reduce the magnitude of this problem, we excluded regular users of aspirin and NSAIDs from the PHS, thus studying only new regular users, and observed no reduction of colorectal cancer risk with aspirin (5) and NSAIDs (3).

Unfortunately, the referenced detailed assessment of the quality of each study (1, 2) was not available from the Agency for Healthcare Research and Quality or the authors. However, the authors mentioned concerns about our aspirin analysis in the PHS because of “contamination by intervention” and a reduced standardized mortality ratio for colorectal cancer compared with that of the U.S. population (1). Randomized aspirin treatment does not threaten the validity of our posttrial study of self-selected aspirin use. Any carryover of putative aspirin effects would bias the results toward a reduced colorectal cancer risk in posttrial aspirin users. Reduced standardized morbidity ratios (5) are ubiquitous in volunteer studies and generally do not bias measures of relative risk.

Finally, because only 3 cohort studies of NSAIDs and colorectal cancer are included in their analysis (2), the failure to include our observational study on NSAIDs and colorectal cancer in the PHS (3) largely reduces the value of the summary estimates. Taken together, we believe that the authors overstate the overall benefits of aspirin (1) and NSAIDs (2) on colorectal cancer. Putting greater weight on randomized evidence available at the time of the review and observational studies with new-user designs would tip the harm to benefit balance even more toward harm for individuals who are at average risk for colorectal cancer. These considerations hold despite the recent additional evidence from randomized trials (6).

Til Stürmer, MD, MPH
Julie E. Buring, ScD
Robert J. Glynn, PhD, ScD
Brigham and Women’s Hospital, Harvard Medical School
Boston, MA 02215

Potential Financial Conflicts of Interest: Dr. Stürmer receives investigator-initiated research funding and support as principal investigator from the National Institute of Aging, National Institutes of Health. Dr. Buring has received investigator-initiated research funding and support as principal investigator from the National Institutes of Health (National Heart, Lung, and Blood Institute; National Cancer Institute; and National Institute of Aging) and Dow Corning Corporation; research support for pills and/or packaging from Bayer HealthCare and the Natural Source Vitamin E Association; and honoraria from Bayer for speaking engagements. She also serves on an external scientific advisory committee for a study by Procter & Gamble. Dr. Glynn receives investigator-initiated research funding and support as principal investigator from the National Institute of Aging and the National Heart, Lung, and Blood Institute, National Institutes of Health. Drs. Stürmer and Glynn do not accept personal compensation from any pharmaceutical company, although they receive salary support from unrestricted research grants from pharmaceutical companies to the Brigham and Women’s Hospital.

References

IN RESPONSE: We want to respond to Dr. Stürmer and colleagues, regarding the systematic review posting on the Agency for Healthcare Research and Quality Web site. The appendices were inadvertently omitted from the Web posting. The mistake has been corrected, and the full systematic review is now available on the Web site at www.ahrq.gov/clinic/uspstf07/aspcol/aspcoloes.pdf (1). We regret the frustration this error has caused.

Mary B. Barton, MD, MPP
Marion M. Torchia
Agency for Healthcare Research and Quality
Rockville, MD 20850
Reversible Hepatic Decompensation in Primary Biliary Cirrhosis Due to Hypercoagulability

Background: Patients with decompensated liver disease due to primary biliary cirrhosis (PBC) rarely recover without liver transplantation.

Objective: To describe 2 sisters with PBC who developed decompensated liver disease associated with a hypercoagulable state (heterozygosity for the factor V Leiden mutation) that resolved with anticoagulation.

Case Report: The first patient was a 47-year-old woman with stage II PBC that was diagnosed in 1996. She was asymptomatic, had detectable antimitochondrial antibodies, and had normal serum bilirubin and albumin levels. Despite treatment with ursodiol and colchicine and trials of methotrexate and then prednisone (for a possible overlap syndrome), her liver disease worsened (1). By 2000, she had developed biopsy-proven cirrhosis, ascites, grade-2 esophageal varices, and lower extremity edema. Her albumin level was 26 g/L, platelet count was 109 cells/L, and prothrombin interna -

g/L, her international normalized ratio was 1.3, and her platelet count was 54 × 10^9 cells/L. Her portal vein was patent. One year later, she developed painful deep venous thrombosis in the left leg. Warfarin was added to therapy. Six months later, her albumin level had increased to 38 g/L. She has now been well for 3 years with warfarin, ursodiol, and colchicine therapy. In February 2007, her albumin level was 38 g/L; alkaline phosphatase level, 98 µkat/L; ALT level, 33 U/L; bilirubin level, 13.7 µmol/L (0.8 mg/dL); and platelet count, 81 × 10^9 cells/L.

Discussion: Both sisters had PBC, heterozygosity for the factor V Leiden mutation, and unusual clinical courses. Both developed cirrhosis and portal hypertension despite aggressive medical therapy and improved remarkably after warfarin treatment was begun. Biochemical tests of liver function normalized, and their portal hypertension lessened. The first patient has no signs of hypersplenism or portal hypertension after 7 years of treatment, whereas the second patient is clearly better after 3 years.

Their clinical courses suggest that intrahepatic thrombosis contributed to their worsening liver disease (2). We hypothesize that chronic inflammation in the liver that is characteristic of PBC triggered low-grade clotting in small branches of the intrahepatic central or portal veins. However, we did not identify thrombosis in the central or portal veins on the liver biopsy specimens, although this may have been because of sampling. Furthermore, heterozygosity for the factor V Leiden mutation confers only a weakly prothrombotic state, leaving uncertainty whether it was the cause of the thrombotic diathesis in these patients. Nevertheless, the remarkable clinical response to anticoagulation suggests that intrahepatic thrombosis should be considered as a potential contributor to hepatic deterioration in patients with PBC, especially those with a personal or family history of thrombotic events.

Hannah M. Lee, MD
Manish Amin, MD
Marshall M. Kaplan, MD
Tufts-New England Medical Center
Boston, MA 02111

Edward F. Herlihy, MD
Hawthorn Medical Group
New Bedford, MA 02747

Potential Financial Conflicts of Interest: None disclosed.

References

CORRECTIONS

Correction: Benzodiazepines and Hip Fractures

In a recent Letter to the Editor on benzodiazepines and hip fracture (1), the last sentence of the second paragraph should have read: “After adjustment for the various risk factors and medication use, sleep
problems were significantly related to falls, but after similar adjustments, the use of psychotropic medication was not.”

Reference

Correction: Serum and Biliary Insulin-like Growth Factor I and Vascular Endothelial Growth Factor in Determining the Cause of Obstructive Cholestasis

The recent article on serum and biliary insulin-like growth factor I and vascular endothelial growth factor in determining the cause of obstructive cholestasis (1) contained an error. In the Editors’ Notes, the final sentence of the Contribution section should have read, “Serum IGF-I levels and biliary VEGF levels did not distinguish accurately among extrahepatic cholangiocarcinoma, benign pancreatic disease, and pancreatic cancer.”

Reference