Prediction Rules Must Be Developed According to Methodological Guidelines

TO THE EDITOR: We read with interest the article by Chan and colleagues (1). They described development of a prediction rule that could be used for pretest probability assessment in pregnant women with suspected deep venous thrombosis (DVT). Prediction rules are valuable tools in daily clinical practice because they provide absolute risks for individual patients, and these risks can be used to determine treatment choices or further diagnostic work-up. However, before a prediction rule can be used safely in daily clinical practice, the methodological soundness of the development steps needs to be assured. Chan and colleagues made some methodological choices that we disagree with.

First, the authors included predictors in the prediction rule if they increased the c-statistic by at least 0.03. However, this approach has been severely criticized, because it is less sensitive than measures based on the likelihood ratio model chi-square or other global measures of model fit (2–4). The authors should have used the likelihood ratio model chi-square to include predictors in the model, if predictor selection is to be used at all (not advised).

Second, the authors did a univariate analysis and entered predictors that were significantly associated with presence of DVT at a P value of 0.05 or less into a multivariate model. Selecting predictors at such a low P value has been strongly advised against, because it leads to poor model performance when evaluated on new patients (external validation), especially in small data sets (5, 6). This is particularly important in this data set because the authors explored too many predictors, considering the low number of events: 11 predictors on 17 events. The authors acknowledged that they needed at least 5 to 10 events per predictor when they developed the rule. They state that, using an initial model with 6 predictors, they needed between 30 and 60 events. However, the authors did not account for the univariate step when making this calculation. Taking this analysis into account, 55 to 110 events would have been needed for the initial 11 predictors that were analyzed.

Third, the authors did not use any imputation method to handle the missing values in their data set. This means that they performed a complete case analysis and analyzed only the data of the patients with complete records. One of the predictors in the model (difference in calf circumference) was missing in 46 of the 194 patients, which means the information of only 75% of the patients was analyzed. Besides a loss of power, it is widely acknowledged that ignoring the missing values in a data set by conducting a complete case analysis may bias study results (7, 8).

Prediction rules are valuable tools for daily clinical practice. Not surprisingly, the number of articles on prediction rules in the medical literature has increased enormously (it more than doubled between 1995 and 2005) (9). Most of the articles concern development of prediction rules. Considering this rapid increase, we believe these prediction rules should be developed according to methodological guidelines. We therefore recommend that researchers developing prediction rules use these guidelines (4–6, 10–13). A selection of the recommended references can also be found in the Information for Authors section of the Annals Web site (14).

Potential Conflicts of Interest: None disclosed.

References
under the receiver-operating characteristic curve), it is unlikely that any other significant predictors were missed by using this method.

We also agree with Dr. Janssen and colleagues’ second point. The low event rate in our study reflects the reality seen in the few diagnostic studies of DVT in pregnant women. The prevalence of DVT in this particular cohort of patients is low (<10%). Despite our best efforts over 7 years, we were limited to 17 events (out of 194 patients). The rule that one should have 5 to 10 events per variable in a multivariable logistic model is based on the fact that having fewer events leads to unstable variable estimates. Thus, it would have been unreasonable to attempt to fit a single model with 11 independent variables to our data, but this does not imply that there is a problem with fitting 11 single-variable models.

The last comment by Dr. Janssen and colleagues is also valid. Our 2 objectives were to determine how clinicians find DVT in pregnant patients by subjective means and whether any “objective” predictors exist that can help clinicians to do this. We believe that we have achieved both objectives. We share all the concerns raised by Dr. Janssen and colleagues, including those regarding the development of prediction rules. We repeatedly emphasized throughout our article that this rule should not be applied in daily practice until it has been properly validated.

In pregnant women, symptoms mimicking DVT are common (for example, leg swelling and pain). At the very least, our study will increase awareness that when certain symptoms (for example, left leg presentation and asymmetry) are present, one should be more vigilant for the presence of DVT and arrange for appropriate testing.

Whee-Shian Chan, MD, MSc
Women’s College Hospital
Toronto, Ontario M5G 1B2, Canada

Tim Ramsay, MSc, PhD
Ottawa Health Research Institute
Ottawa, Ontario K1Y 4E9, Canada

Jeffrey S. Ginsberg, MD
McMaster University Medical Centre
Hamilton, Ontario L8N 3Z5, Canada

Potential Conflicts of Interest: None disclosed.

The Net Clinical Benefit of Warfarin Anticoagulation in Atrial Fibrillation

TO THE EDITOR: With regard to the article by Singer and colleagues (1), calculating net clinical benefit of warfarin therapy in atrial fibrillation is an important concept (2). However, 3 other issues should be considered in the future.

First, warfarin therapy affects other major events and may substantially reduce myocardial infarction and death (3).

Second, because control of the international normalized ratio (INR) varies in different settings and because a 10% improvement in the amount of time the INR is in the therapeutic range (TTTR) has been associated with a greater than 10% reduction in event rates (3, 4), the relationship between INR control and event rates must be more thoroughly defined before it can be determined how INR control may affect the net clinical benefit in one’s own setting. Veeger and colleagues (5) found that the bottom quartile of approximately 4000 patients was in range only 10% to 20% of the time and accounted for more than half of the major events. Jones and colleagues (4) found that the bottom quartile of 2223 patients was in range only about 28% of the time. Furthermore, because event rates increase exponentially as the INR moves further out of range, being slightly out of the target range may have little effect, whereas being at the extremes (INR, <1.5 or >5.0) may carry a very high risk. Therefore, one must know what the event rates were when the INR was in the target range TTR ≥ 0.3 INR units, below an INR of 1.5, and above an INR of 5.0 to estimate the net clinical benefit in one’s own setting.

Third, one must consider how evolving methods to improve INR control may alter the net clinical benefit. As noted by Hart and Halperin (2), the TTR of 65% reported by Singer and colleagues is often considered “high.” By combining INR self-testing and computer management, however, Harper and Pollock (6) reported an 80% TTR with no INRs greater than 5. An interim analysis of our similar study (7) found a TTR of 78.9%, which increased to 94% when the range was expanded slightly by ±0.3 INR units. Approximate TTR values were 90% for the top quartile and 60% for the lowest quartile, and the percentage of time that the INR was greater than 5 or less than 1.5 was only 0.27%. Such improved INR control is estimated to yield a 30% to 50% reduction in both thromboembolic and major bleeding events compared with “typical” management—changes that should have a substantial effect on the net clinical benefit of warfarin.

Henry Bussy, PharmD
The University of Texas at Austin
Austin, TX 78712

Potential Conflicts of Interest: Consultancies: Genesis Advanced Technologies. Stock ownership or options (other than mutual funds): Inverness Medical. Patents received/pending: 3 vitamin K antagonist products.

References
7. Bussey HI, Walker MB, Bussey-Smith KL, Fei CR. Interim analysis of triple

TO THE EDITOR: Current guidelines from the American College of Cardiology/American Heart Association recommend that warfarin be given to patients with atrial fibrillation on the basis of stroke risk. This strategy maximizes stroke reduction only if the effect of warfarin is uniform across all risk factors. Singer and colleagues (1) present data that suggest otherwise. Their subgroup analysis found that men and women with atrial fibrillation have markedly different relative risk reductions when treated with warfarin (adjusted relative risk reduction, about 34% in men vs. 55% in women). Formal testing for treatment by covariate interaction (not reported by the authors) might confirm that the benefit of warfarin varies substantially in different subgroups of patients. Although female sex is not currently considered an indication for warfarin, the data suggest that certain patient subgroups may reap greater benefit from warfarin therapy than others. The presence of hypertension supports the use of warfarin according to current guidelines, but subgroup analysis shows little, if any, additional benefit in patients with hypertension compared with those without hypertension (absolute stroke reduction with warfarin, 1.11 per 100 person-years in patients with hypertension vs. 0.95 per 100 person-years in patients without hypertension). An optimum strategy for stroke reduction in patients with atrial fibrillation should be based not on factors that imply a high risk for stroke (as recommended by current guidelines), but rather on factors that imply a greater benefit of treatment.

Vikram Budhraja, MD
Lincoln Medical Center
Fayetteville, TN 37334

Potential Conflicts of Interest: None disclosed.

Reference

IN RESPONSE: We support Dr. Bussey’s emphasis on the importance of maximizing TTR. The benefits and risks of warfarin therapy in atrial fibrillation are strongly dependent on maintaining the INR in the optimum range (1, 2). That is why we reported the overall TTR, 65%, for our ATRIA (Anticoagulation and Risk Factors In Atrial Fibrillation) cohort. At lower TTR values, the expected net clinical benefit will be lower, and at higher TTR values, the net clinical benefit will increase regardless of clinical risk category as long as the distribution of out-of-range values does not change markedly. As pointed out, having more extreme out-of-range low or high INR values will have a strong negative effect on the net benefit conferred by any given level of TTR. It is remarkable that warfarin’s striking efficacy has been shown in studies in which the INR was in range only about two thirds of the time (for example, in the BAFTA [Birmingham Atrial Fibrillation Treatment of the Aged Study] trial [3]). Substantially higher overall TTRs can be achieved in selected populations (2). More widespread use of organized anticoagulation management services, INR self-testing, improved dosing algorithms, and possibly genetic testing for warfarin sensitivity all have the potential to effect more widespread improvement in TTRs and to increase net clinical benefit in patients with atrial fibrillation.

The final sentence in Dr. Budhraja’s letter aptly summarizes the motivation for our article. To generate an optimum estimate of the adjusted absolute net benefit of warfarin therapy, we allowed the effect of warfarin to vary by patient subgroup in our models, independent of the statistical significance of any individual interaction term. In fact, most of the interaction terms were not statistically significant and, except for the lowest stroke-risk categories, the interaction effects were small. Before firm conclusions that the relative risk reduction conferred by warfarin differs in any specific subgroup are drawn, such effect modification should be confirmed in other large databases, especially those from randomized trials.

As Dr. Budhraja notes, ischemic stroke risk off warfarin is only one determinant of net clinical benefit, but it is a very important one. Unfortunately, current risk stratification schemes for patients with atrial fibrillation have very limited accuracy (4). Future research should emphasize improved risk prediction in atrial fibrillation. Better risk prediction will facilitate achieving increased net benefit of warfarin at both the population and individual patient levels.

Daniel E. Singer, MD
Yuchiao Chang, PhD
Massachusetts General Hospital
Boston, MA 02114

Alan S. Go, MD
Kaiser Permanente of Northern California
Oakland, CA 94612-2304


References
Comparing Costs and Quality of Care at Retail Clinics With Those of Other Medical Settings

TO THE EDITOR: I read the article by Mehrotra and colleagues (1) with great interest; however, I do have some concerns. First, the study compares episodes in patients who were first seen in different clinical settings (retail clinic, physician office, urgent care clinic, and emergency department). Patients may not have received care in the same clinical setting throughout the episode. Although the authors mention that the percentages of episodes with any follow-up visits were similar for retail clinics, urgent care centers, and physician offices and higher for emergency departments, we do not have information on the number of follow-up visits or the clinical settings in which they were conducted. Several patients who were first seen in retail clinics may have received follow-up in urgent care clinics or physician offices. Therefore, the quality-of-care measures and preventive services attributed to retail clinics could have been due to follow-up in other clinical settings.

Second, the study participants were matched by diagnosis codes but not by severity of symptoms, leading to a high likelihood of selection bias. It is possible that patients who initially visited a retail clinic had less-severe illness, and most may not have required laboratory testing or prescription medications. The overall cost for these patients may have been even lower if they were first seen in an urgent care clinic or a physician’s office, because the higher cost of evaluation in these settings would be balanced by the money saved on unnecessary laboratory testing and medications.

Finally, the authors matched the study participants by income category but not by race or geographic location. Retail clinics are more likely to be present in areas with higher median income and lower population percentage of black persons and are less likely to exist in medically underserved areas (2). Therefore, the population served by the retail clinics may differ significantly from that served by urgent care clinics and physician offices, even though all patients had the same health plan.

Rajan Kochar, MD, MPH
University of Texas Health Science Center at Houston
Houston, TX 77030

Potential Conflicts of Interest: None disclosed.

References

TO THE EDITOR: The outcomes measured by Mehrotra and colleagues (1) show the fundamental weaknesses of retail clinics. Otitis media and pharyngitis are self-limited conditions whose clinical courses are not substantially altered by medical interventions in most cases. Many physicians in other developed countries do not commonly prescribe antibiotics for otitis media (2, 3) or pharyngitis (4).

Many retail clinics display a menu of diagnoses with prices. If a patient truly has one of these conditions, Mehrotra and colleagues successfully showed that the quality of care provided by retail clinics for these conditions is probably reasonable. I have heard stories from persons who have visited these clinics who tell me that no matter what the presenting set of circumstances, the patient will usually be diagnosed with one of the conditions on the menu.

Drawing conclusions on the basis of diagnosis codes from the International Classification of Diseases, Ninth Edition, misses the point. The quality issue is how an undifferentiated symptom, such as lower abdominal pain, is managed, not how a urinary tract infection is treated. If primary care clinics in the United States used the International Classification of Primary Care codes, such research could probably occur without having to look through thousands of charts (5).

Ultimately, retail clinics will probably neither help nor hurt U.S. health care in any measurable way. With 75% of health care expenditures attributable to chronic diseases (6), the method of treating a self-limited sore throat is of little consequence.

Richard A. Young, MD
John Peter Smith Family Medicine Residency Program
Fort Worth, TX 76104

Potential Conflicts of Interest: None disclosed.

References
expert opinion and some clinical trial evidence support longer treat-
m ent durations). Little comfort is provided by knowing that retail
clinics are not doing any worse than traditional care settings for
management of urinary tract infections when the traditional settings
are doing so poorly; on the contrary, the latter finding is quite
concerning.

James R. Johnson, MD
Veterans Affairs Medical Center
Minneapolis, MN 55417

Potential Conflicts of Interest: Consulancies: Ortho-McNeil. Grants
received: Bayer, Wyeth, Procter & Gamble, Merck.

References
paring costs and quality of care at retail clinics with that of other medical settings for 3
2. Warren JW, Abrutyn E, Hibel JR, Johnson JR, Schaeffer AJ, Stamm WE. Guid-
elines for antimicrobial treatment of uncomplicated acute bacterial cystitis and acute
pyelonephritis in women. Infectious Disease Society of America (IDSA). Clin Infect

IN RESPONSE: Dr. Kochar is correct that the follow-up care of an
episode at a retail clinic could occur at another care site. But we
doubt that this affected our quality scores. Fewer than 20% of epi-
sodes included any follow-up visits. Also, our quality measures gen-
erally focused on care around the first visit. For example, we looked
at antibiotic prescriptions for otitis media that were filled on the day
of the first visit or in the subsequent 2 days.

We shared Dr. Kochar’s concern that patients who presented to
a physician’s office could be more ill and that this might drive some
of the cost differences we observed. Yet, our sensitivity analyses that
directly addressed this issue did not support this concern. The major
driver of cost differences is reimbursement for the first visit. Costs
that are probably related to severity of illness (for example, labora-
 tory costs and follow-up visits) were less important. Nonetheless,
as we note in our article, residual differences could exist
between the patient populations at the care sites, although we feel
that matching on income, level of illness, and insurance plan
minimizes these differences.

Although we agree with Dr. Young that most cases of otitis
media and pharyngitis are self-limited, 11.4% of all pediatric primary
care visits are for just these 2 problems (1). Seeking care for these
problems is the established norm in our society, and it is unclear
whether this is a fair criticism of retail clinics. We agree that retail
clinics are not a magic bullet for reduction of health care costs. In
our own models, we estimate that $2 billion in cost savings would
result if retail clinics become widespread, which constitutes less than
0.1% of health care spending.

We disagree with Dr. Young that the more important quality
issue is how undifferentiated symptoms, such as lower abdominal
pain, are managed. Proper management of uncomplicated urinary
tract infections is an important issue and is the subject of much
research and several guidelines (2). Also, our goal was to compare the
care at retail clinics with that of other care sites, and retail clinics do
not manage undifferentiated abdominal pain.

Consistent with our previous work (3), our current study does
find problems with quality across all the care sites. We agree with Dr.
Johnson that efforts to improve quality of care across the health care
system are critical.

Ateev Mehrotra, MD
University of Pittsburgh School of Medicine
Pittsburgh, PA 15213

N. Marcus Thygeson, MD
HealthPartners
Bloomington, MN 55425

Elizabeth A. McGlynn, PhD
RAND
Santa Monica, CA 90407

Potential Conflicts of Interest: None disclosed.

References
care physicians, and emergency departments: a comparison of patients’ visits. Health
348:2635-45. [PMID: 12826639]

CLINICAL OBSERVATION

Metformin Stinks, Literally

Background: Metformin causes adverse gastrointestinal effects,
including diarrhea, nausea and vomiting, flatulence, distention, and
abdominal pain, and these effects often necessitate discontinuing the
drug (1). Occasionally, taste distortion has been reported, but this
event seems to be infrequent (1). Adverse gastrointestinal effects
seem to be less common with extended-release metformin (2).

Objective: To describe a characteristic of metformin that has not
been reported previously in the medical literature.

Case Reports: Case 1 was an adult man with diabetes and severe
visual impairment. He had taken brand-name metformin (Gluco-
ephage, Bristol-Myers Squibb, New York, New York) for several years
without reporting adverse effects. He was switched to an immediate-
release, generic version of metformin, which he discontinued. He
reported that it smelled like “dead fish” and nauseated him. He was
rechallenged with an extended-release, generic formulation of met-
formin. On follow-up, he reported no further problems and was able
to continue the medication without ill effect.

Case 2 was an adult man with diabetes, not visually impaired,
who discontinued generic metformin because the odor of the drug,
which he described as “fishy,” nauseated him. The patient was un-
willing to try an extended-release formulation.

Our cases meet the criteria for adverse drug reaction as defined
by Edwards and Aronson (3). Our cases can be subclassified as ad-
verse reaction, non–dose-related (bizarre) (3).

Discussion: We searched PubMed under the terms metformin,
side effects, adverse reactions, smell, and odor and found no case reports
or studies discussing the odor of metformin or listing odor as a
potential adverse drug reaction. In fact, one standard drug reference (4) states that metformin is odorless.

Although reaction to the odor of metformin has not been reported in the medical literature, hundreds of postings to message boards on the Internet (5–8) note the peculiar odor of the drug, which is also well known to pharmacists. In an informal survey of several pharmacists (including the 2 pharmacist coauthors), all agreed that they could readily identify metformin by the odor, which was variously described as fishy or “like old locker-room sweat socks.” The odor reportedly varies considerably between generic versions and seems to be less noticeable with film-coated (extended-release) formulations. Although Brittain (4) states metformin is odorless, Bristol-Myers Squibb, the manufacturer of Glucophage, is aware of complaints about odor (Liang J, May JR. Personal communication.).

We wonder why this reaction to metformin has not been previously reported. Patients may report that metformin nauseates them but do not further elaborate or distinguish this as a visceral reaction to the smell of the medication. Physicians probably interpret the report of nausea in light of well-known gastrointestinal side effects of the medication and do not pursue the issue further.

Our cases show that the distinctive odor of metformin (independent of other, well-known gastrointestinal adverse effects of the medication) causes patients to stop taking the drug. The effect seems to be certain, but idiosyncratic and patient dependent. This is a nonpharmacologic adverse effect of the drug.

When patients stop taking metformin, physicians should consider inquiring more closely about revulsion to the odor of the medication. Trial of a film-coated, extended-release formulation may be a reasonable approach in such cases. Further research is needed to determine whether odor revulsion is a common but heretofore unrecognized reason for discontinuation of metformin.

Allen L. Pelletier, MD
Ann M. Butler, RPh, MD
Ralph A. Gillies, PhD
Medical College of Georgia
Augusta, GA 30912-3500

J. Russell May, PharmD
The University of Georgia College of Pharmacy
Athens, GA 30602-2351

Potential Conflicts of Interest: None disclosed.

References

Correction

Correction: Hypomagnesemia Induced by Several Proton-Pump Inhibitors

In the Clinical Observation letter by Broeren and colleagues (1), there are errors in the axis labels of the Figure. On the bottom graph, the left-hand y-axis label should say “Magnesium, mmol/L,” and the right-hand y-axis label should say “Magnesium Excretion, mmol/d.” These have been corrected in the online version.

Reference