

*Editor's Note:* As this month's ASCP Corner offering, we are pleased to present an article adapted from the first annual Donald L. Klein Lifetime Research Award Lecture, given by Dr A. John Rush at the 2014 ASCP Annual Meeting.



## Narrowing the Gaps Between What We Know and What We Do in Psychiatry

A. John Rush, MD<sup>a,\*</sup>

**T**his report addresses the gaps between what we know, need to know, and do as practitioners. These research-practice gaps leave much clinical decision-making to experience rather than evidence, which can directly impact patient outcomes. Historically, the challenge of applying group data to individual patients and the importance of evidence to inform clinical decision-making have not been widely recognized. How and when to select, combine, or sequence treatments remain more art than science. Intellectual silos, academic incentives, and publication practices contribute to these gaps. In this article, case examples are used to illustrate these contentions.

Suggestions to reduce these gaps are proposed. Studying more clinically representative populations and framing research questions to better inform clinical decision-making are essential. Using innovative, clinically relevant trial designs—such as point-of-care trials—can reduce cost, increase generalizability, and hasten clinical acceptance of research findings. Development of clinical decision rules combined with regular reporting of likelihood ratios can help bridge the gap between group-based research data and its application to individuals. Greater attention to reporting results in clinically informative terms, conducting secondary analyses to inform clinical decision-making, and reducing the burden of patient-oriented research are discussed and illustrated with case examples.

The current climate of limited resources for clinical care and research may best be viewed as a stimulus for researchers to develop more impactful, less expensive, and more informative research that fills knowledge gaps and can change clinical practice to improve outcomes.

This report examines the nature, importance, and reasons for the gaps between what we know from research, what we need to know, and what we do in the clinic: research-practice gaps. It also makes suggestions for how to narrow these gaps. Patients, families, clinicians, care system managers, regulators, health care policy makers, and payers are all stakeholders in patient-oriented research. Different types of patient-oriented research address their diverse interests. For example, regulators and industry sponsors focus largely on efficacy and safety issues for new medications and devices via phase 1–3 studies. Clinicians and patients want to know how, when, for whom, and under what conditions both new and current treatments are best provided—questions not addressed by trials conducted for regulatory approval. Managers and policy makers understandably want to know about tradeoffs

and costs, especially the value added (or not) by new, potentially more expensive treatments. This report will focus on patient and clinician perspectives, which substantially overlap.

Figure 1A summarizes common clinical tasks such as screening, diagnosing, treatment targeting, and monitoring outcomes. Treatment targeting entails the selection of the best treatment for specific patients or, conversely, the avoidance of treatments that are rather certain to fail for specific patients (the core of personalized medicine).

### WHY ARE THERE RESEARCH-PRACTICE GAPS?

Both conceptual and operational factors contribute to these gaps. Conceptual challenges include historical factors; scientific and academic incentives; intellectual silos among clinicians, researchers, and other stakeholders; and the failure to adequately recognize and address the tension entailed in nomothetic and ideographic perspectives in the development, reporting, and application of evidence. Operational challenges include budgets, funding priorities, costs, publication practices, and the growing research burdens placed on both patients and researchers.

#### Conceptual Factors

**Historical assumptions.** We have generally assumed that mental health, medical, and surgical practitioners, once armed with a new clinical test or treatment, will figure out by clinical experience how to best use it. Clinical experience would result in consequent wisdom (ie, the “art of medicine”), so research would not be needed.

Once, the surprisingly wide practice variations were recognized. The potential benefit of reducing ineffective or unsafe variations was clear. Practice guidelines were developed to summarize the evidence and recommend evidence-based practices. These recommendations, while helpful, have been of limited value due to the dearth of empirical evidence regarding who, how, when, and in what contexts any specific treatment or clinical test should be selected, avoided, placed in a sequence, combined, or otherwise personalized. This evidence is often just not available.

**Scientific and academic incentives.** Historically, academic incentives have been highly focused on individual researchers, but this focus also inadvertently interferes with the development, management, and recognition of members of multidisciplinary scientific teams. These teams are essential in the design and execution of successful patient-oriented research. Team members include clinicians from multiple medical disciplines who have extensive disease and treatment knowledge, statisticians with often diverse expertise, economists, and laboratory-based collaborators (geneticists, imagers, immunologists, etc). As part of this individual focus, academic incentives have strongly emphasized authorship order and independent idea generation, which too often incentivize

<sup>a</sup>Duke-National University of Singapore, Singapore

\*Corresponding author: A. John Rush, MD, 6, The Academia, 20 College Road, Singapore 169856 (john.rush@duke-nus.edu.sg).

*J Clin Psychiatry* 2015;76(10):1366–1372

dx.doi.org/10.4088/JCP.15ac10309

© Copyright 2015 Physicians Postgraduate Press, Inc.

secrecy rather than collaboration and potentially more creative science.

In brief, overvaluing individual competition may undermine achieving valuable synergies. While individuals remain critical for innovation, the science model is no longer embodied by an isolated genius achieving great creations. Collaboration has become essential, as no one person can master all disciplines well. Other fields have figured out how to retain the emphasis on individuals while recognizing and incentivizing team science. Ours must as well.

**Intellectual silos.** Intellectual silos, a natural consequence of specialization, have developed among and between clinicians, researchers, and other stakeholders. Academic institutions have created such divisions as tracks, department subunits, and subspecialization. While sensible on the face of it, an unintended consequence of these silos is that each group is largely uninformed regarding what the others do. This lack of awareness with regard to existing mutual interests and potential benefits reduces the chances—albeit unintentionally—of developing innovative research that crosses field boundaries.

These intellectual silos have also resulted in markedly diminished collaboration between scientists in academia and industry. The consequences have been reduced funding from industry, greater engagement by industry of scientists abroad, and reduced scientific interactions.

**Ideographic and nomothetic perspectives.**<sup>1</sup> Practicing clinicians typically shift between 2 distinct intellectual perspectives: a nomothetic, or group-based, perspective typically provided by scientific studies and an ideographic, or individual-patient-focused, perspective. Basically, clinicians have to assess and manage 1 patient at a time. To help the individual patient, the clinician must properly apply, adapt, or even sometimes ignore the available scientific evidence provided in the form of group-based information. Researchers study samples or groups, from which they try to glean general principles and understandings. The clinicians bridge the nomothetic (group) and individual perspectives.

What is true for the group may not be true for the individual, especially when the groups being studied are highly heterogeneous with regard to having therapeutic or adverse effects with any treatment or test.

Over the last 4 decades, evidence-based medicine has aimed to fill this gap. Its limitations are significant, as the information that forms the basis for evidence-based medicine is often not applicable to many patients who are seen in practice. Furthermore, studies that develop evidence to inform the above clinical tasks are few and far between. That is, studies that demonstrate the efficacy of the treatment strategies, such as medications, rely largely on regulatory trials. While such studies are not uncommon, studies that define the tactics of how, when, for whom, and under which contexts to use these strategies are relatively rare.

For example, most blinded controlled trials in depression are conducted by industry for regulatory purposes. These trials logically limit participants to those with little general medical or psychiatric comorbidity, treatment resistance, chronicity, or suicidality, all of which limit generalizability.<sup>2</sup> Actually, three-quarters of Sequenced Treatment Alternatives to Relieve Depression (STAR\*D) enrollees (representative depressed outpatients who regularly receive antidepressant medication in practice) would have been ineligible for any regulatory trial due to these factors. Furthermore, those with more general medical and psychiatric comorbidities have lower likelihoods of response and remission<sup>3</sup> and greater risk of relapse in follow-up.<sup>4</sup>

**Table 1. Outperformance of Longer Depression Measures by Their Shorter Versions<sup>a</sup>**

| Measure             | % Improvement | % Respond | % Remit |
|---------------------|---------------|-----------|---------|
| HDRS <sub>17</sub>  | 42.1          | 45.7      | 36.1    |
| HDRS <sub>6</sub>   | 46.9          | 50.4      | 44.9    |
| <i>P</i> value      | < .001        | < .001    | < .001  |
| IDS-C <sub>30</sub> | 42.1          | 45.8      | 32.7    |
| IDS-C <sub>6</sub>  | 47.0          | 52.3      | 43.4    |
| <i>P</i> value      | < .001        | < .001    | < .001  |

<sup>a</sup>Based on Ostergaard et al.<sup>6</sup>

Abbreviations: HDRS<sub>6</sub>=6-item Hamilton Depression Rating Scale, HDRS<sub>17</sub>=17-item Hamilton Depression Rating Scale, IDS-C<sub>6</sub>=6-item Inventory of Depressive Symptomatology—Clinician-Rated, IDS-C<sub>30</sub>=30-item Inventory of Depressive Symptomatology—Clinician-Rated.

Thus, when a medication comes to market, we know we have a safe and effective new treatment but only in a very limited, largely unrepresentative population. Evidence-based medicine and consequently practice guidelines rely heavily on controlled trials conducted for regulatory purposes. While these trials provide the highest level of certainty, results may or may not apply to most of our patients. Consequently, the Institute of Medicine and others have emphasized the need to develop learning health care systems to systematically accumulate experiences and evidence that inform the clinical tasks.<sup>5</sup>

#### Operational Factors

Several operational factors contribute to the research-practice gap. Three are illustrated herein.

**Research burdens and costs.** The requirements, time, staff burdens, and costs of conducting patient-oriented research have grown dramatically over the years. The goals—greater validity, quality, and safety—are laudable, but have they been realized? Has the greater cost produced value?

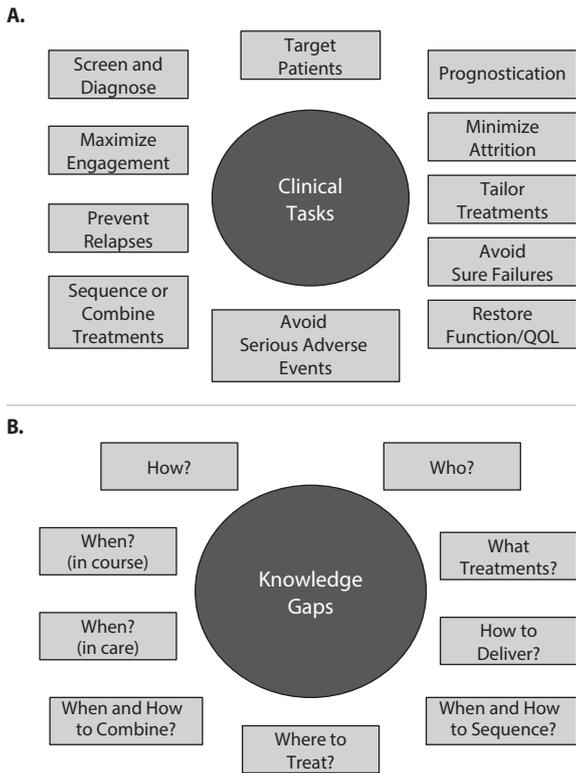
Research, an inherently risky business, cannot be made risk-free. We need to strike a reasonable balance between “as safe as feasible” and fictional perfection. Every layer of oversight increases time and cost and increases burdens on patients, providers, and researcher staffs. Fewer studies are funded when each costs more.

Procedural burdens can be reduced. For example, consent forms are now so long that they are actually *less* informative. The STAR\*D consent form exceeded 10 pages! Why not use a 1-page executive summary in all consent forms with added pages for details? There are only 5 or 6 key items that inform participants (eg, study purpose, expected risks and benefits, compensation for injury, expected participant obligations, right to discontinue). One page should suffice and enhance participants’ understanding. Streamlining of contract and intellectual property processes, as well as centralized institutional review boards that meet several times a week, can reduce cost and save time while preserving quality, especially in multisite trials.

Simplification of research outcomes would reduce cost and save time. Table 1, adapted from Ostergaard et al,<sup>6</sup> shows that the 6-item Hamilton Depression Rating Scale and the 6-item Inventory of Depressive Symptomatology—Clinician-Rated each outperform their longer versions. Briefer simplified research outcome measures also facilitate their use in practice.

**Funding priorities.** Research to address the above-noted critical clinical tasks has generally been less highly regarded scientifically than research that discovers new disease mechanisms, treatment targets, or therapeutic molecules. Patients, clinicians, and payers, however, disagree with this prioritization. Evidence that informs

**Figure 1. (A) Common Clinical Tasks and (B) Current Knowledge Gaps That Affect Clinical Decision-Making**



Abbreviation: QOL = quality of life.

these tasks can make the difference between a successful or failed treatment or test. Nevertheless, this prioritization affects resource allocation and thus limits research support that could reduce the gap.

In fact, until recently, funding for this arena of research has been largely ignored. The Federal government has begun to address these concerns by the creation of the Patient-Centered Outcomes Research Institute (PCORI).<sup>7</sup> The annual PCORI budget of about \$500 million is extremely modest given its mission of addressing all medical and surgical conditions. Through 2014, PCORI spent 16.3% of its budget on cancer, 10.3% on mental/behavioral health and another 10.6% on trauma and injury, and 7.6% on cardiovascular disease (Grayson Norquist, MD; e-mail communication; 2015), which suggests that the need for this type of research is substantial.

Industry, on the other hand, does not support truly definitive phase 4 studies, especially those that search for populations who might not benefit from the product or who could otherwise define when the product should not be used.

While research that amplifies our understanding of brain function or leads to new potentially innovative biologics is an obvious priority, few would argue that it should be an exclusive priority. Certainly, the National Cancer Institute and National Heart, Lung, and Blood Institute support both basic and more patient-oriented research. Learning how to better use our current interventions despite our limited comprehension as to mechanisms and target engagement remains a critical task by which to inform clinical decision-making with evidence that can change both the quality of practice and patient outcomes.

While some rebalancing of the National Institute of Mental Health portfolio could be very helpful, we need to be innovative and

cost-conscious in creating new funding mechanisms for patient-oriented research. For example, industry and health care systems, perhaps in combination with PCORI, could co-fund registries for the first 25,000 to 100,000 people who receive a new biologic that could be provided at reduced cost as an incentive for participation in research. Such efforts might identify who should not receive the treatment and clarify its side effects in a wider, more representative patient sample.

**Publication practices.** Two publication practices contribute to the research-practice gap: overvaluing the *P* value (and a positive study) and the low priority for publishing replication studies.

Many patient-oriented research reports do not initially specify what is to be considered a clinically meaningful or actionable finding. They do report the power and effect size that the study can detect. Of course, a statistically significant effect may not be clinically meaningful. Also, clinicians may have neither time nor inclination to ponder the practical implications of often complex research results.

The importance of replication studies and their rarity in psychiatry were highlighted recently by Kapur et al.<sup>8</sup> Replications are critical when patient samples are highly heterogeneous, as in psychiatry. Replications do not make headlines, boost the journal's impact factor, or garner high praise from promotion and tenure committees as much as those exciting initial reports. However, replications are actually more important than initial reports because they can either establish or dismiss a "fact." Finally, the chances that a replication study proposal will actually be funded are very low. National Institutes of Health (NIH) review committees would view such a proposal as lacking in innovation. Yet, without replication, how do we ever establish facts?

**ARE THE RESEARCH-PRACTICE GAPS IMPORTANT?**

Figure 1B summarizes some of our clinical decision-making knowledge gaps. These gaps lead to wide variation among practitioners. Some variations may be excellent, but others may be detrimental to achieving good outcomes. These gaps are due in part to the absence of studies in relevant populations.

These knowledge gaps are not unique to our fields. Physicians "currently struggle to apply new medical knowledge to their own patients, since most evidence regarding the effectiveness of medical innovations has been generated by studies involving patients who differ from their own and who were treated in highly controlled research environments."<sup>9(p2161)</sup>

To illustrate the importance of these sorts of knowledge gaps, consider the fact that greater numbers of general medical conditions (GMCs) are associated with lower rates of symptomatic remission with citalopram.<sup>3</sup> Patients with more GMCs are often excluded from registration trials, yet depression is a risk factor for developing GMCs and vice versa. With no placebo-controlled trials to determine whether antidepressants are effective in these patients with substantial numbers of GMCs, we could actually be doing harm and not know it! A similar case can be made for depressed pregnant women, for whom we have no placebo-controlled randomized trials to determine whether antidepressants are effective. Certainly these sorts of knowledge gaps—and there are many—are important to patients and clinicians.

**WAYS TO BRIDGE THE RESEARCH-PRACTICE GAPS**

We provide several suggestions for how we as researchers might bridge these gaps.

It is illegal to post this copyrighted PDF on any website.

### 1. Study More Relevant Populations

As noted above, our practice guidelines rely too heavily on evidence from studies of patients who are not representative of the majority of patients whom we treat in practice. Studies of representative populations would be more immediately applicable in general practice.

This can be accomplished through the use of registries, learning health care systems, and innovative designs (see below). Health care systems not only provide platforms for clinically informative research but also can benefit directly from the research discoveries. As part of that effort, we need to implement designs with lower patient and researcher burden that appeal to more patients while retaining internal validity and, whenever possible, randomization. Patients need to be educated about the benefits—for themselves, their families, and their friends—of becoming research participants. Effective broad patient engagement ensures optimal generalizability and reduces cost and time.

### 2. Ask More Clinically Informative Questions

Research questions are too often framed with little clinician input, perhaps due to the intellectual silos. Consider the differences between clinician- and researcher-framed questions. Researchers may want to know whether a specific biomarker is significantly associated with a particular descriptive diagnosis or a particular treatment outcome. Practitioners share this interest, but they also really want to know whether the association is clinically meaningful enough, which is rarely specified a priori or reported.

Another too-familiar example is that researchers largely focus on symptoms, while patients and practitioners care as much or more about function and quality of life, which typically change more slowly than symptoms.<sup>10</sup> These longer-term, critically important outcomes could be clarified either by registries or by extension of follow-up periods after short-term trials.

As another example, consider the STAR\*D trial, which asked whether bupropion or bupropion sustained release (SR) was a more effective augmenting agent with citalopram. The remission rates did not differ by either the 17-item Hamilton Depression Rating Scale or the 16-item Quick Inventory of Depressive Symptomatology.<sup>11</sup> Case closed? Not quite.

Clinicians want to know: which is the better treatment to use? To answer this question, efficacy, side effects, and quality of life (QOL) are all key outcomes. Using a modified intent-to-treat sample, bupropion-SR was associated with fewer side effects, greater reduction in depressive symptoms, and better QOL (trend level).<sup>12</sup> One could even make the argument for a descriptive look at the completers who benefited enough to stay on the drug to determine their side effect burden, symptom status, and QOL, along with their course over the following 3–6 months! All of this information would be clinically informative, though not definitive given the attrition rates and potential biases.

### 3. Use Clinically Informative Trial Designs

Over the last 2 decades, many new designs have been developed that could be used to address our knowledge gaps in clinical decision-making. They include pragmatic clinical trials,<sup>13</sup> point-of-care (POC) trials,<sup>14</sup> adaptive design trials,<sup>15</sup> adaptive treatment trials (sequential, multiple assignment, randomized trial [SMART] design),<sup>16</sup> and registries.

Pragmatic clinical trials may differ from registration trials on the basis of at least 10 parameters (Table 2).<sup>17</sup> One, several, or even all of these parameters can be controlled or allowed to vary by design. When all parameters are uncontrolled, the design is comparing 2

**Table 2. Ten Parameters That May Differentiate Practical Clinical Trials From Registration Trials<sup>a</sup>**

| Parameter                                  |
|--|
| Eligibility criteria                       |
| Type of outcomes                           |
| Follow-up intensity                        |
| Primary analysis                           |
| Participant compliance                     |
| Practitioner adherence                     |
| Practitioner expertise (comparison)        |
| Practitioner expertise (experimental)      |
| Flexibility of intervention (experimental) |
| Flexibility of intervention (comparison)   |

<sup>a</sup>Based on Thorpe et al.<sup>17</sup>

treatments, for example, under routine care conditions (as in some comparative efficacy studies). When all parameters are controlled, the design mimics a registration trial, with high internal validity but very limited external validity.

STAR\*D was a hybrid efficacy-effectiveness trial that tightly controlled some elements, such as the delivery of medication, but not others that enabled representation of typical practice, such as patient and provider selection and follow-up. Which parameters to control depends on the question(s) to be asked.

Point-of-care trials can reduce cost and yet generate results that are immediately applicable in practice. Fiore et al<sup>14</sup> designed a trial to determine which of 2 methods is better for defining the proper insulin dose in newly diagnosed diabetic inpatients: a sliding scale or weight-based regimen. This trial requires *only* participant consent. The trial does not affect management of any of the patients' other medical conditions, nor does it require any additional diagnostic procedures, extra follow-up clinical or research visits, or any data collection beyond the electronic medical record that provides the primary outcome (length of hospital stay) and secondary outcomes (hemoglobin A<sub>1c</sub> and readmission rates within 30 days of discharge).

Point-of-care trials can be used to address many knowledge gaps in clinical decision-making (eg, how, for whom, or when a particular treatment or laboratory test is clinically valuable). Since electronic medical records are used in POC trials, the trials can also estimate costs associated with the various options being studied. Finally, unlike registration trials, POC trials can change clinical guidelines, and therefore practice, because they can use randomization to develop Level 1 evidence in generalizable patient samples. They are of immediate value to system managers and policy makers as well.

### 4. Develop and Test Clinical Decision Rules

Clinical decision rules (CDRs)<sup>18,19</sup> or clinical prediction rules<sup>20,21</sup> are being developed to provide a transparent, evidence-based approach to identifying and weighing requisite information to accomplish the clinical tasks noted above in individual cases. CDRs may be formed using clinical findings alone or include basic laboratory information. CDRs are in wide use in emergency medicine<sup>22–26</sup> and elsewhere in medicine and surgery,<sup>27,28</sup> but not in psychiatry, though they should be of value because they provide predictions before action is taken.

CDRs can estimate the likelihood of a particular diagnosis, the potential value of a laboratory test, or beneficial or adverse effects of a particular treatment<sup>28,29</sup> or program element.<sup>23,24,27</sup> CDRs often provide clinicians with a range of probabilities as to the outcome.<sup>30–32</sup> Naturally, the clinical value and validity of these predictions vary depending on patient and contextual issues.<sup>33</sup>

To illustrate, consider the CDR that estimates the likelihood of a deep venous thrombosis (DVT).<sup>31,34</sup> The clinician completes the 9 items that indicate the presence or absence of various common clinical findings or risk factors such as local tenderness, leg swelling, presence of an active cancer, and collateral superficial veins. The scale ranges from 0 to 8, with higher scores associated with a greater likelihood of DVT. Specific thresholds or cutoffs are recommended for unlikely, moderately likely, and so on. The CDR is applicable to each patient. In addition, the CDR provides a uniform metric across clinicians to estimate risk. It can also be used to estimate other outcomes such as response to treatment or adverse effects.

The Hunter criteria set, one of the few CDRs in psychiatry, estimates the likelihood of serotonin syndrome with sensitivity and specificity of 84% and 97%, respectively.<sup>35</sup> This CDR is based on just 9 key clinical findings (spontaneous, inducible and ocular clonus, agitation, diaphoresis, tremor, hyperreflexia, temperature, and hypertonicity/hyperrigidity). Typical of CDRs, it simplifies in a clinician-friendly fashion the efforts needed to arrive at the diagnosis by focusing on a limited number of clinical findings.

When CDRs are combined with likelihood ratios (see below), CDRs offer the possibility of personalizing treatment choices, bridging the nomothetic-ideographic divide, and establishing a consistent clinically based platform upon which we can evaluate potentially informative laboratory tests that are now being developed.

### 5. Report Results in Terms That Better Inform Clinicians and Other Stakeholders

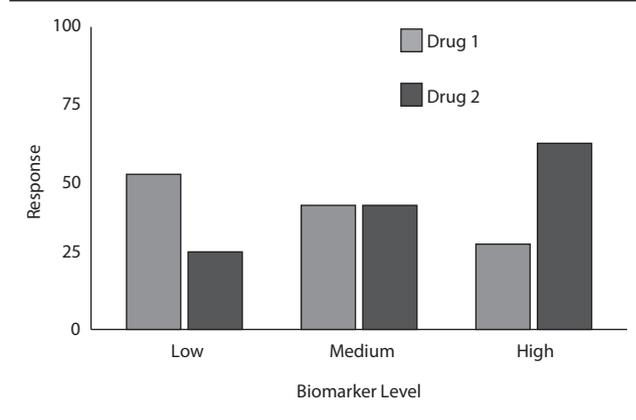
The research-practice gaps can be reduced by making research reports more clinically informative. For example, while researchers focus largely on statistical significance, clinicians also want to know whether the effect size reported is clinically meaningful, as well as generalizable. In comparative trials, the number needed to treat (NNT) or to harm is now more commonly reported. Each estimates the number of people who must receive the treatment to produce a positive or negative outcome in 1 patient. If the NNT is large between 2 active treatments, then the choice of treatment may rest on factors other than efficacy (eg, risk of side effects, cost, convenience).

Biomarkers provide a different challenge. Consider a hypothetical biomarker (Figure 2) that appears to be a moderator of outcome with 2 medications. That is, a higher value is associated with greater chances of response with drug 2 and lower chances with drug 1. The converse also pertains. For patients in the middle range of the biomarker, both agents look acceptable. With adequate sample size, statistical significance of the moderator effect seems very likely.

Clinicians, however, need more information to consider applying this intriguing, statistically significant finding in the clinic. What thresholds define high and low? How many patients in routine practice fall into the low, medium, and high ranges? Even better, what are the sensitivities, specificities, predictive values, and likelihood ratios that are associated with each possible biomarker level for each drug?

As with most research findings, clinicians need to apply these group data to individuals with different situations. Even if the biomarker predicts a lower likelihood of response, clinicians may use the drug anyway if there are few other options or if the patient is especially fearful of side effects associated with the other drug. Clinically useful information is gleaned even if the test does not recommend one drug over the other (ie, the level is in the middle range), because these patients can choose between the 2 drugs

**Figure 2. Hypothetical Predictive/Prognostic Markers: What Is a Meaningful Difference?**



on the basis of considerations other than efficacy, such as side effects, cost, and convenience. The biomarker provides a basis for shared decision-making between patient and clinician. In brief, in addition to evidence of statistical significance, the actual values and their performance are critical to clinicians' understanding and application of research findings. A second suggestion is to specify a priori and subsequently actually report what are considered clinically meaningful outcomes. This would help clinicians more readily grasp patient-oriented research findings.

The difference between a clinically meaningful and clinically actionable finding in patient-oriented research deserves mention. A clinically meaningful difference suggests that outcomes would be improved for a meaningful proportion of patients (say at least 10%)—akin to number needed to treat.<sup>36</sup> Of course, the threshold chosen to define clinically meaningful depends on contexts such as risks, costs, and availability of other options.

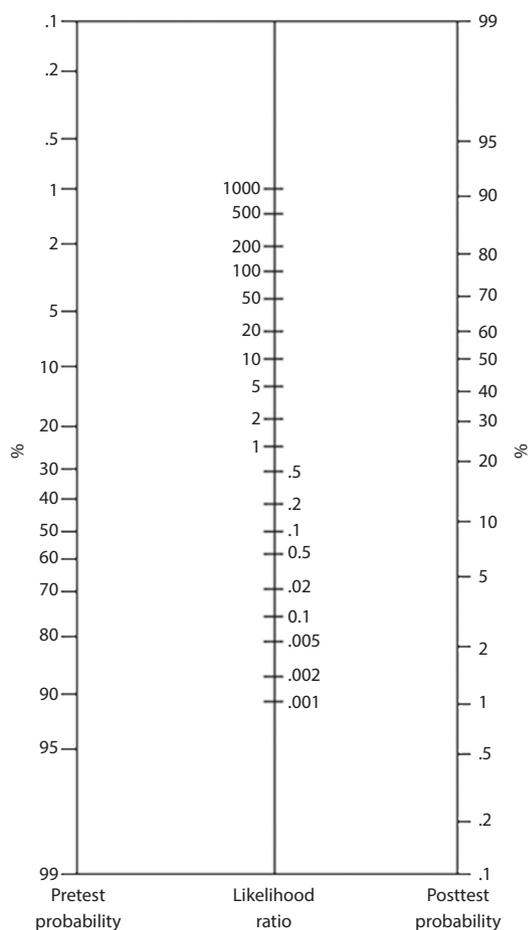
*Clinically actionable* implies a relatively greater degree of certainty that a meaningful proportion of patients would be rather certain to benefit if action were taken based on some aspect of the research findings. In the above biomarker example, all 3 groups defined by their biomarker levels would derive clinically meaningful information from knowing their biomarker levels. Clinical action would be more likely as the biomarker levels approach the extreme levels in either direction, although context would still affect whether and when to act.

A third suggestion to help clinicians convert group-based findings into individually informative, potentially actionable information is to report likelihood ratios (LRs).<sup>37,38</sup> Sensitivity, specificity, and positive and negative predictive values reflect population characteristics, so they may not translate reliably to individual patients. LRs provide a practical way to make sense of diagnostic tests and prognostic biomarkers, including those that inform treatment selection. A large positive likelihood ratio (+LR) and a small negative likelihood ratio (-LR) indicate that the test is clinically useful:

$$\text{Positive LR} = \text{sensitivity}/(1 - \text{specificity})$$

$$\text{Negative LR} = (1 - \text{sensitivity})/\text{specificity}$$

LRs do not vary across different populations because they are ratios. LRs help clinicians use group data generated for scientific reports to make personally tailored patient estimates. LRs quantitate the likelihood of the outcome under study, such as diagnosis,

**Figure 3. The Bayes Nomogram<sup>a</sup>**

<sup>a</sup>Based on Page and Attia.<sup>39</sup> Pretest probability is located on the first axis and joined to the appropriate likelihood ratio on the second axis. The posttest probability is then read off the third axis.

response, relapse, adverse events, or the value of doing an added new laboratory test. LRs are intuitive; larger positive LRs indicate a higher likelihood and smaller positive LRs a lower likelihood of the positive outcome. Consider a 65-year-old man with a positive stress test (sensitivity = 65%, specificity = 89%). The positive LR is 5.9, and the negative LR is 0.39. The likelihood of coronary artery disease is increased 6-fold with a positive test.

While LRs are quite informative and applicable across populations, we can further personalize this information if we combine the test performance expressed as LRs with an estimation of the pretest probability of the condition. We can estimate the pretest probability by clinical intuition/experience (using history, physical examination, laboratory tests, family history, etc) or by CDRs as noted above. To combine LRs and pretest probabilities, we use the Bayes nomogram (Figure 3)<sup>39</sup> that combines the pretest probability and the test characteristics expressed as LRs to provide a posttest estimate for that individual patient!

To illustrate, consider the above 65-year-old with a positive stress test with an LR+ of 5.9. Let us assume that he has a pretest likelihood of coronary artery disease of 20% (estimated clinically). Locate 20% in the left column of Figure 3. Connect that point to the middle column at 5.9 and continue drawing the line to the right hand column, which is at about 62%. If his pretest probability was

10%, his posttest probability is about 40%. These estimates provide clinical information that is valid across groups, and they allow clinicians and patients to participate in decision-making.

This personalized approach can be applied to new laboratory tests as they develop; it can also be used to compare different tests and evaluate test sequences. This approach can also be used to predict other outcomes such as treatment responses, side effects, relapses, and remissions. This opportunity also makes a strong case for developing CDRs, whenever possible, to better estimate pretest probabilities.

## 6. Conduct Secondary Analyses and Data Mining to Inform Clinical Decision-Making

Secondary analyses are cost-efficient ways to develop hypotheses, and with appropriate safeguards for multiple testing, they can even test hypotheses that have been framed a priori from other datasets. As large databases are developed, such analyses will likely prove quite useful as they directly apply to “real world” patients.

A host of secondary analyses using STAR\*D clinical and genetic data have led to hypotheses that were subsequently tested in new studies.

As an additional illustration, consider the discontinuation trial by Weihs and colleagues<sup>40</sup> that found that, in recurrent major depressive disorder, continuing on bupropion treatment resulted in lower relapse rates (about 30%) than switching to placebo (about 50%). A secondary analysis could have asked an additional clinically important question: Which patients can stop taking bupropion (ie, switch to placebo) and *not* relapse?

Funding for such analyses should be made easily and rapidly available with small grants and just-in-time peer review, from either PCORI or NIH.

## SUMMARY

There are many important research-practice gaps. Evidence to inform clinical decision-making is remarkably scant. A greater research focus on more clinically informative patient samples, clinically relevant questions, the use of innovative trial designs and analyses, and reporting of results in clinically understandable and potentially actionable ways can help reduce these gaps. Efforts to reduce intellectual silos, lessen the burden of research, and revise some academic incentives and publication practices would also be of help in reducing these gaps and hopefully improving patient care and outcomes.

**Potential conflicts of interest:** Dr Rush has received consulting fees from Brain Resource, Eli Lilly, Emmes Corporation, Lundbeck A/S, MedAvante, National Institute on Drug Abuse, Santium, Stanford University, Takeda; speaking fees from the University of California at San Diego, Penn State Hershey Medical Center, New York State Psychiatric Institute, and the American Society for Clinical Psychopharmacology; royalties from Guilford Publications and the University of Texas Southwestern Medical Center; a travel grant from CINP; and research support from Duke–National University of Singapore. Through the University of Texas Southwestern Medical Center, he has a potential financial interest in the Inventory of Depressive Symptomatology and several variations of it. Dr Gelenberg has reviewed this article and found no evidence of influence from these relationships.

**Funding/support:** None reported.

**Previous presentation:** Parts of this article were presented at the first annual Donald L. Klein Lifetime Research Award Lecture at the American Society of Clinical Psychopharmacology Annual Meeting, June 2014, Hollywood, Florida.

**Acknowledgments:** Dr Rush acknowledges the editorial support of Jon Kilner, MS, MA (Pittsburgh, Pennsylvania), funded by the author.

## REFERENCES

- Ideographic versus nomothetic approaches. In: Marshall G. *A Dictionary of Sociology*. 2nd ed. New York, NY: Oxford University Press; 1998.
- Wisniewski SR, Rush AJ, Nierenberg AA, et al. Can phase III trial results of antidepressant medications be generalized to clinical practice? a STAR\*D report. *Am J Psychiatry*. 2009;166(5):599–607.
- Trivedi MH, Rush AJ, Wisniewski SR, et al; STAR\*D Study Team. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR\*D: implications for clinical practice. *Am J Psychiatry*. 2006;163(1):28–40.
- Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR\*D report. *Am J Psychiatry*. 2006;163(11):1905–1917.
- Olsen LA, Aisner D, McGinnis JM; IOM Roundtable on Evidence-Based Medicine. *The Learning Healthcare System: Workshop Summary*. Washington, DC: National Academies Press; 2007.
- Ostergaard SD, Bech P, Trivedi MH, et al. Brief, unidimensional melancholia rating scales are highly sensitive to the effect of citalopram and may have biological validity: implications for the Research Domain Criteria (RDoC). *J Affect Disord*. 2014;163:18–24.
- Selby JV, Lipstein SH. PCORI at 3 years—progress, lessons, and plans. *N Engl J Med*. 2014;370(7):592–595.
- Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry*. 2012;17(12):1174–1179.
- Schneeweiss S. Learning from big health care data. *N Engl J Med*. 2014;370(23):2161–2163.
- Mintz J, Mintz LI, Arruda MJ, et al. Treatments of depression and the functional capacity to work: corrections. *Arch Gen Psychiatry*. 1993;50(3):241.
- Trivedi MH, Fava M, Wisniewski SR, et al; STAR\*D Study Team. Medication augmentation after the failure of SSRIs for depression. *N Engl J Med*. 2006;354(12):1243–1252.
- Bech P, Fava M, Trivedi MH, et al. Outcomes on the pharmacopsychometric triangle in bupropion-SR vs buspirone augmentation of citalopram in the STAR\*D trial. *Acta Psychiatr Scand*. 2012;125(4):342–348.
- Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA*. 2003;290(12):1624–1632.
- Fiore LD, Brophy M, Ferguson RE, et al. A point-of-care clinical trial comparing insulin administered using a sliding scale versus a weight-based regimen. *Clin Trials*. 2011;8(2):183–195.
- Chow SC, Chang M. *Adaptive Design Methods in Clinical Trials*. 2nd ed. Boca Raton, FL: CRC Press; 2011.
- Murphy SA, Oslin DW, Rush AJ, et al; MCATS. Methodological challenges in constructing effective treatment sequences for chronic psychiatric disorders [published online November 8, 2006]. *Neuropsychopharmacology*. 2007;32(2):257–262.
- Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol*. 2009;62(5):464–475.
- McGinn TG, Guyatt GH, Wyer PC, et al; Evidence-Based Medicine Working Group. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. *JAMA*. 2000;284(1):79–84.
- Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Ann Emerg Med*. 1999;33(4):437–447.
- Laupacis A, Sekar N, Stiell IG. Clinical prediction rules: a review and suggested modifications of methodological standards. *JAMA*. 1997;277(6):488–494.
- Wasson JH, Sox HC, Neff RK, et al. Clinical prediction rules: applications and methodological standards. *N Engl J Med*. 1985;313(13):793–799.
- Harnan SE, Pickering A, Pandor A, et al. Clinical decision rules for adults with minor head injury: a systematic review. *J Trauma*. 2011;71(1):245–251.
- Stiell IG, Greenberg GH, McKnight RD, et al. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Ann Emerg Med*. 1992;21(4):384–390.
- Stiell IG, Greenberg GH, McKnight RD, et al. Decision rules for the use of radiography in acute ankle injuries: refinement and prospective validation. *JAMA*. 1993;269(9):1127–1132.
- Stiell IG, Wells GA, Hoag RH, et al. Implementation of the Ottawa Knee Rule for the use of radiography in acute knee injuries. *JAMA*. 1997;278(23):2075–2079.
- Stiell IG, Wells GA, Vandemheen K, et al. The Canadian CT Head Rule for patients with minor head injury. *Lancet*. 2001;357(9266):1391–1396.
- Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med*. 1997;336(4):243–250.
- Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med*. 2004;141(12):920–928.
- Landefeld CS, Goldman L. Major bleeding in outpatients treated with warfarin: incidence and prediction by factors known at the start of outpatient therapy. *Am J Med*. 1989;87(2):144–152.
- Hebert JJ, Fritz JM. Clinical decision rules, spinal pain classification and prediction of treatment outcome: a discussion of recent reports in the rehabilitation literature. *Chiropr Man Therap*. 2012;20(1):19.
- Riddle DL, Wells PS. Diagnosis of lower-extremity deep vein thrombosis in outpatients. *Phys Ther*. 2004;84(8):729–735.
- Riddle DL, Hoppener MR, Kraaijenhagen RA, et al. Preliminary validation of clinical assessment for deep vein thrombosis in orthopaedic outpatients. *Clin Orthop Relat Res*. 2005;432(4):252–257.
- Newman DH, Schriger DL. Rethinking testing for pulmonary embolism: less is more. *Ann Emerg Med*. 2011;57(6):622–627, e3.
- Wells PS, Owen C, Doucette S, et al. Does this patient have deep vein thrombosis? *JAMA*. 2006;295(2):199–207.
- Dunkley EJ, Isbister GK, Sibbritt D, et al. The Hunter Serotonin Toxicity Criteria: simple and accurate diagnostic decision rules for serotonin toxicity. *QJM*. 2003;96(9):635–642.
- Chatellier G, Zapletal E, Lemaitre D, et al. The number needed to treat: a clinically useful nomogram in its proper context. *BMJ*. 1996;312(7028):426–429.
- Harrell FE Jr, Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. *JAMA*. 1982;247(18):2543–2546.
- Reid MC, Lane DA, Feinstein AR. Academic calculations versus clinical judgments: practicing physicians' use of quantitative measures of test accuracy. *Am J Med*. 1998;104(4):374–380.
- Page J, Attia J. Using Bayes' nomogram to help interpret odds ratios. *Evid Based Med*. 2003;8:132–134.
- Weihl KL, Houser TL, Batey SR, et al. Continuation phase treatment with bupropion SR effectively decreases the risk for relapse of depression. *Biol Psychiatry*. 2002;51(9):753–761.

ASCP Corner offerings are not peer reviewed by the Journal but are peer reviewed by ASCP. The information contained herein represents the opinion of the author.

Visit the Society Web site at  
[www.ascpp.org](http://www.ascpp.org)