Learning by Doing:

Can Our Collective Experiences as Clinicians Improve Mental Health Care?

A. John Rush, MD, and Tony Tramontin, PhD

Decision-Making in Current Mental Health Care

In the delivery of mental health care, 3 stakeholder groups (clinicians/program heads, managers/administrators, and payors) are making decisions based more on impression and experience than on evidence—akin to the tale of the 3 blind mice. Clinicians and program heads base their decisions on personal experiences augmented by published studies that too often include patients and treatment delivery methods dissimilar to their own. Managers and administrators reactively allocate resources, with insufficient information about clinically relevant outcomes. Payors know what they spend but often do not know how outcomes are impacted. These challenges obstruct the development of a "learning health care system," as envisioned by the Institute of Medicine, where outcomes are systematically obtained, compiled, and analyzed to inform and revise clinical and programmatic decisions.[1]

Can Clinicians Learn From Each Others' Experiences?

At the heart of mental health care lie myriad clinical decisions that significantly impact patient outcomes, program effectiveness, and the cost of care. Currently, individual clinicians learn from their day-to-day interactions with their own patients as they make diagnostic, theranostic (treatment-related), and prognostic decisions at virtually every visit. Clinicians' abilities to learn from each other's experiences, however, remain largely untapped, save for infrequent, time-consuming requests for second opinions.

If clinical and diagnostic features (eg, symptoms, diagnoses, life circumstances/stresses, history of illness), treatment characteristics, clinical outcomes (eg, symptoms, function, quality of life [QoL]), and service utilization outcomes (eg, hospitalizations, emergency department use, medication, therapy costs) for each patient were readily available, compiled, and analyzed, the decisions by clinicians, administrators, and payors would be far more evidence-based than is presently the case. This compiled information would facilitate the automatic sharing of insights, enabling essentially thousands of second opinions in real-time, for each patient at each visit. Simultaneously, administrative and payor decisions about whether and where to invest for better outcomes, safer care, or better patient engagement would also become more evidence based.

For instance, the early identification of nearly half the patients who drop out of buprenorphine treatment for opioid use disorder within the first 6 months could help clinicians and program heads develop, target, and evaluate interventions designed to assist these patients.[2] Alternatively, identifying patients at highest risk for hospitalization in the ensuing 6 months would help to target interventions to reduce this potentially preventable worsening.[3]

Obstacles to "Learning by Doing"

While the above-noted potential clinical, administrative, and economic value of such a learning health care system has been widely discussed, the obstacles to realizing these aspirations are substantial. These obstacles derive largely from 3 critical gaps: an information (data collection) gap, a knowledge (evidence) gap, and a learning gap.[4,5]

The Information (Data Collection) Gap. The information gap refers to the dearth of systematically collected, recorded, and analysis-ready information that pertains to diagnostic, treatment, and clinical outcomes.

Diagnoses are typically made by unstructured clinical interviews conducted by mental health professionals with widely varied backgrounds, despite evidence that structured interviews are more reliable.[6] While new more reliable and efficient diagnostic approaches have been developed, such as computerized adaptive testing (CAT),[7,8] entry into practice has been gradual. CAT or artificial intelligence administered structured interviews or could readily increase access and reliability, save time, improve cost efficiency, and provide an important second opinion when treatments fail or diagnoses are uncertain.

Treatment—in terms of the doses and types of medications and psychotherapy attendance—is easily found in the electronic health record (EHR), but adherence to the prescribed treatments remains unknown. Yet, adherence remains a major care delivery challenge.[9] In addition, missed appointments (both the intentionally cancelled and "no shows") are difficult to cull from many EHRs, yet their value in predicting relapse as well as their impact on the cost of care is important. Revision in the EHR to make this information readily available is called for.

Turning to treatment and longer-term prognostic outcomes, clinicians routinely assess overall disease severity, function, side effects, and

often QoL at nearly every treatment visit to decide what to do or not do. While this information is typically found in clinical notes, it is not systematically recorded with an agreed upon metric. Diagnostic and service utilization information is available in claims data, but these data have not been readily tied to treatment processes and outcomes found in the EHR. Without outcomes, the collective learning potential from clinicians' day-to-day experiences remains unrealized.

Our reluctance to record outcomes has historical roots. Mental health and general medical care were provided by individuals trained in the apprenticeship model, wherein learning came from 1 patient at a time under senior guidance. Training focused on developing clinical skills and change processes (eg, gaining insight, solving marital, or self-esteem issues) rather than on what some saw as more "superficial" outcomes such as symptoms or daily function. As clinical trials matured, however, measurement of syndrome/symptom severity, function, and QoL developed to obtain regulatory approval, though these measurements rarely entered daily practice—indeed, they were not designed for easy use in practice. With the recent recognition that measurement-based care enhances outcomes and the growing emphasis on value-based health care,[10] efforts to define a "common set of measures" for mental health practitioners and clinical researchers are underway.[11] Some quality improvement efforts are now requiring measures from time to time in clinical practice.[12] The development of artificial intelligence will further push us to be diligent but efficient information providers.[5]

Bridging the Information Gap. To bridge the information gap, we need to (1) decide which outcome domains to assess and record in the EHR; (2) decide on the tool(s) to assess each domain; and (3) evaluate whether these assessments are useful to the 3 stakeholder groups.

We would suggest 1–3 domains (symptom/syndrome severity, daily function, and QoL). These domains are germane to all diagnoses and treatments including medication, psychotherapy, and brain simulation treatments. The rationale is that (1) these domains inform clinical decisions (2) they are valued by patients; (3) each has prognostic value, yet the combination of all 3 is optimal[13]; and (4) single item or short itemized self-reports are in the public domain and applicable to virtually all psychiatric diagnoses. Simple measures could be completed by clinicians (eg, the Clinical Global Index of Severity[14]) or by patients (eg, either the Patient Global Index of Severity[15] or the Visual Analog Scale[16]). These simple ordinal scales (eg, none, mild, moderate, and severe) are likely sufficient for daily clinical decision-making and have a built-in minimal clinically significant change metric (eg, change from moderate to mild). Regularly obtained simpler measures may yield greater predictive precision than more lengthy measures obtained less frequently (see, for example, Taquet et al[3]). Longer itemized scales can be used when a programmatic or clinical problem requires greater precision. They can also be used to benchmark ordinal scales, facilitating the translation of clinical research results based on the longer scales into practice.[17] However, the prospective evaluation of whether mandatory simple global outcome acquisitions at each visit are sufficiently accurate to be of value to clinical or program decision-makers is essential.

In the future, the further application of natural language processing to clinical notes may offer a scalable solution to obtain outcomes or other clinical information. Additionally, capturing voice features reflecting mood, energy, impulsiveness, and other states, or remotely monitoring activity to assess activation, might provide additional data and granularity. These innovative approaches can further bridge the information gap.

The Knowledge (Evidence) Gap. Information is not evidence. Evidence (knowledge) is generated when information is used to answer specific diagnostic, theranostic, prognostic, or cost-effectiveness questions. Do we actually need evidence based on real-world information? Don't we have sufficient knowledge from clinical practice guidelines? Maybe and maybe not. If your patient is fairly similar to patients in the trials that formed the basis for the guidelines, then real-world data may add little additional knowledge. But most of our patients have treatment histories, comorbid medical or psychiatric conditions, psychosocial circumstances, or other factors unlike registration trial participants, such that outcomes found in trials may not apply to many of our patients.[18] Here, real-world data are needed and informative.

In addition, we need real-world evidence to address questions that trial data do not. For example, what is the efficacy or safety of a treatment when used off-label or when combined with other medications? When during a multistep treatment sequence is a particular treatment (eg, TMS) effective? Only real-world data can address these questions. Tactical treatment decisions also require real-world data (eg, what starting doses and dose escalation rates are useful and safe in a 75-year-old person with diabetes and other medical conditions?). Real-world evidence is needed to address patient and program management issues for which trial data are largely unavailable. For example, when, how, and for whom should psychotherapy be stopped? What is the relapse rate when medications are stopped as opposed to being continued? Real-world evidence is also needed to establish the costs and benefits of various programs, staffing, or administrative policies, as well as specific clinical practices. Understanding the short- and long-term costs or savings resulting from specific interventions is crucial for making informed decisions.

These evidence gaps are numerous and evolve over time, as do our

treatments, their indications, and practice patterns. Establishing an accessible, easily queried living library of past and current experiences can partially address these evolving questions.

The Learning Gap. The compilation and analysis of real-world data provide more direct and precise observational evidence that suggests possible courses of action across diverse patients, treatments, and clinical contexts. This enhancement in evidence quality can significantly impact decision-making. However, the possible courses of action suggested by these observational data must be tested prospectively to determine whether these suggestions (which are basically inferences from the observations) actually produce the desired effects and at what cost and risk.

For example, if a specific antidepressant augmenting agent seems to be more effective than another, one could query the data to determine whether the patients benefiting from each are similar or distinct. But to directly test whether there is indeed differential effectiveness, randomization is needed. If clinicians and patients are in equipoise, randomization with patient consent could be considered in a particular program.

As another example, an intervention that seems to be associated with higher treatment dropout or poorer adherence rates might be identified. Selected clinics or practitioners could decide to minimize the use of this intervention, while others do not. The prospective results between clinics (or practitioners) could determine whether the reduced use of the apparently less desirable medication actually improves retention or adherence.

In addition, the compilation of data across many patients enhances our ability to identify, anticipate, and hopefully avoid untoward side effects or other undesired outcomes. For instance, evidence as to which patients with dementia are at particularly high risk for cardiovascular arrhythmias when treated with a specific psychotropic medication can be generated. The effects of proscribing that medication on those specific patients can be assessed prospectively. Is the assumed benefit realized? In addition, the impact on the management of those patients who are now not using the medication must be examined.

One final challenge—the dissemination and implementation of these learnings in practice—deserves mention. The implementation of practice guidelines has been uneven, in part because the randomized controlled trials upon which they are based often do not reflect the patients being seen in practice. By addressing the information, evidence, and learning gaps with "real-world" patients from practicing clinicians, we might expect that acceptance and implementation of evidence-based practice changes based on these data would be facilitated.

Just as important in the application of this information is the need to engage patients (and supportive family members when appropriate) in the process of shared decision-making.[19] As the information and evidence gaps are narrowed with data and evidence generated from real-world practice, patients and their clinicians still must collaborate to understand and weigh the risks and benefits to choose among treatment options and sequences. The regular use of simple ratings of disease severity or function as suggested to address the information gap also provides a metric for a conversation about how well or poorly the treatment is going, thereby facilitating shared decision-making.

Conclusion

In summary, clinicians, administrators, and payors make decisions with only modest evidence. Simple outcomes, often patient reported, could facilitate evidence-based decision-making by clinicians, administrators, and payors, thereby providing the foundation for a learning health care system.

References

1. Institute of Medicine. Grossmann C, Powers B, McGinnis JM, eds. *Digital Infrastructure for the Learning Health System: The Foundation for Continuous Improvement in Health and Health Care: Workshop Series Summary*. National Academies Press, National Academy of Sciences; 2011.
2. Ker S, Hsu J, Balani A, et al. Factors that affect patient attrition in buprenorphine treatment for opioid use disorder: a retrospective real-world study using electronic health records. *Neuropsychiatr Dis Treat*. 2021;17:3229–3244.
3. Taquet M, Griffiths K, Palmer EOC, et al. Early trajectory of clinical global impression as a transdiagnostic predictor of psychiatric hospitalisation: a retrospective cohort study. *Lancet Psychiatry*. 2023;10(5):334–341.
4. Friedman C, Rigby M. Conceptualising and creating a global learning health system. *Int J Med Inf*. 2013; 82(4):e63–e71.
5. Kraus B, Sampathgiri K, Mittal VA. Accurate machine learning prediction in psychiatry needs the right kind of information. *JAMA Psychiatry*. 2024; 81(1):11–12.
6. *Handbook of Assessment and Treatment Planning for Psychological Disorders*. 3rd ed. Guilford Publications; 2020.
7. Gibbons R, Wang P. The science of psychiatric measurement. *Psychiatr Ann*. 2023;53(9):400–404.

8. Gibbons RD, Weiss DJ, Frank E, et al. Computerized adaptive diagnosis and testing of mental health disorders. *Annu Rev Clin Psychol*. 2016;12:83–104.

9. Semahegn A, Torpey K, Manu A, et al. Psychotropic medication non-adherence and its associated factors among patients with major psychiatric disorders: a systematic review and meta-analysis. *Syst Rev*. 2020; 9(1):17.

10. Scott K, Lewis CC. Using measurement-based care to enhance any treatment. *Cogn Behav Pract*. 2015; 22(1):49–59.

11. Hermann RC, Palmer RH. Common ground: a framework for selecting core quality measures for mental health and substance abuse care. *Psychiatr Serv*. 2002;53(3):281–287.

12. The Joint Commission. What are the key elements when utilizing a standardized tool to monitor progress towards treatment goals? Accessed March 27, 2024. https://www.jointcommission.org/standards/standard-faqs/behavioral-health/care-treatment-and-services-cts/000002332/

13. IsHak WW, Mirocha J, Christensen S, et al. Patient-reported outcomes of quality of life, functioning, and depressive symptom severity in major depressive disorder comorbid with panic disorder before and after SSRI treatment in the star*d trial. *Depress Anxiety*. 2014;31(8):707–716.

14. Guy W. *Clinical Global Impression. ECDEU Assessment Manual for Sychopharmacology*. National Institute of Mental Health; 1976.

15. Mohebbi M, Dodd S, Dean OM, et al. Patient centric measures for a patient centric era: agreement and convergent between ratings on the Patient Global Impression of Improvement (PGI-I) Scale and the Clinical Global Impressions – improvement (CGI-S) Scale in bipolar and major depressive disorder. *Eur Psychiatry*. 2018;53:17–22.

16. Polomano RC, Galloway KT, Kent ML, et al. Psychometric testing of the Defense and Veterans Pain Rating Scale (DVPRS): a new pain scale for military population. *Pain Med*. 2016;17(8):1505–1519.

17. Rush AJ. Narrowing the gaps between what we know and what we do in psychiatry. *J Clin Psychiatry*. 2015; 76(10):1366–1372.

18. 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.

19. Zisman-Ilani Y, Lysaker PH, Hasson-Ohayon I. Shared risk taking: shared decision making in serious mental illness. *Psychiatr Serv*. 2021;72(4):461–463.

Scan Now
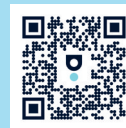
Cite and Share this article at Psychiatrist.com