

Big Science Research Collaboration for Evidence-Based Care

Victor M. Montori, MD, MSc

The Central Question of Medicine

Staring at landscapes made of pill containers amidst mountains of paperwork, patients contemplate their troubling present and their uncertain future. What is best for me? What is best for my family?

At least 3 challenges await those seeking an answer to this central question of medicine. The first challenge results from the paucity of trustworthy comparative effectiveness research that directly addresses patient dilemmas. The second challenge relates to what each person values and would consider best for their own situation. Fitting treatments within the context of each person's life is the third challenge.

Ongoing treatments, to work, have to fit into people's daily routines, weaving seamlessly and constantly into their day-to-day activities.¹ It is hard to predict how a new intervention will interact with the existing care plan, and how the care plan will work with each day's plan. Clinicians cannot discover patient values, preferences, and contexts without interacting meaningfully with patients. To find what is best, clinicians must partner with patients to think, talk, and feel through their situation, use the best research evidence, and craft together the best treatment plan while minimally disrupting their lives.^{2,3}

What sort of evidence enterprise can support that work? Here, I propose that the evidence necessary to support patient-centered care requires a fundamental change in the culture of research. Finding what is best for patients requires generous collaboration for Big Science.

How Little Science Fails Patients

The available research produces evidence that clinicians and patients struggle to use to determine what is best for each patient. It is as if the research enterprise was not fit for this purpose. Consider trials determining the value of a treatment by measuring its relative impact on a primary end point. This end point is often a surrogate marker, a laboratory measurement devoid of patient experience, importance or meaning.⁴ Or it is a composite end point comprised important outcomes (eg, mobility and mortality) combined with less important ones. These trivial ones tend to capture most of the effect and cloud the interpretation of trial results.⁵ These end points are needed to run the briefest and smallest trial, usually with much <1000 very high-risk participants. This choice is then argued beautifully in a sample

size calculation, researchers' closest foray into writing fiction. The budget for this Lilliput project would come, somehow, only a few cents short of the maximum fundable budget. As they efficiently succeed in identifying a winning intervention, the narrow and skewed definition of success in these studies reduces their value to usual patients and their clinicians. Should a patient use a treatment because it can impact a laboratory marker? How to know the extent to which patients like this would be better off—for example, live longer, with less disability or feel better—with this treatment? For patients taking multiple medications already, how to appreciate the incremental value vis-à-vis the incremental burden of treatment?⁶ Are these trials able to provide patients and clinicians with salient reasons to support their decisions, for example, the relative effect of alternatives on outcomes that people can experience and value?

Many large drug and device trials are designed primarily to secure the approval of the US Food and Drug Administration. Approval hinges on successful comparisons against placebo controls rather than against another active option that patients may use instead. But in the frontlines of care, in deciding what is best for our patient, what we need are comparisons among the best, most sensible alternatives. Instead, placebo-controlled trials forces reliance on unreliable indirect comparisons, contrasting each agent's effect against placebo.

A recent development permits the comparison of alternatives using all available studies, including trials that directly compared agents against each other and placebo-controlled trials.^{7,8} Placed in a network of comparisons, we can analyze the scarce head-to-head direct comparisons and the more numerous indirect comparisons to better estimate the relative impact of each agent. Pooling published evidence, however, cannot overcome the limitations in the underlying evidence. Because some was produced to position drugs as market and sales leaders, only partial and biased data sets and trial results are published and available for pooling.^{9,10}

We can hardly compensate in clinical care for inadequate or unreliable evidence. Consider efforts to bring the evidence about antidepressants to bear in helping primary care clinicians and their patients select an antidepressant. Combining what could be used of the published record with the expertise of clinicians and patients, a tool to support this process was

Adapted from a Keynote Address at the First Patient-Centered Outcome Research Institute (PCORI) Annual Meeting, Washington, DC, 2015. Dr Montori is grateful for the generous contribution to the ideas reflected here of his colleagues at the Knowledge and Encounter Research Unit at Mayo Clinic. The ideas expressed here do not necessarily represent views held by PCORI.

From the Department of Medicine, Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, MN.

Correspondence to Victor M. Montori, MD, MSc, Department of Medicine, Knowledge and Evaluation Research Unit Mayo Clinic, 200 First St SW, Rochester, MN 55905. E-mail montori.victor@mayo.edu

(*Circ Cardiovasc Qual Outcomes*. 2016;9:688-692. DOI: 10.1161/CIRCOUTCOMES.116.003358.)

© 2016 American Heart Association, Inc.

Circ Cardiovasc Qual Outcomes is available at <http://circoutcomes.ahajournals.org>

DOI: 10.1161/CIRCOUTCOMES.116.003358

produced.^{11,12} The tool compares available agents on their ability to affect patient weight, sexual function, and sleep quality; on their effects with drug discontinuation; and on their daily use and out-of-pocket cost. One domain is crucially absent from this tool because the evidence of any difference across agents was too unreliable: their impact on depression. The work and expense in conducting trials of efficacy that were not designed, conducted, and published to meet the decision needs of patients and clinicians has supported fantastically effective marketing campaigns. Yet, this evidence cannot be used confidently to figure out what is best for our patients.

Big Data Is Not Great Data

The sexiest solution proposed to address the problem of inadequate evidence for clinical decision making is to harness the power of the so-called Big Data. Big Data makes use of the unintentional and heterogeneous data byproducts of the provision of clinical care and of billing payers for this care. Mining these extremely large observational and administrative data sets requires sophisticated policies, informatics, statistics, and data processing to render the data useful and its use ethical.¹³ Many features of Big Data make it attractive. Investigators can propose myriad projects against the same data set, lowering the cost per question. Questions made feasible by mega-data sets explore the effects of treatments, including those use infrequently, on patient subgroups, and on rare outcomes, on harms, or on outcomes that take a long time to appear. Because these data result from actual practice, they are less affected by the constraints typical of efficacy trials.¹⁴

Unfortunately, in my view, these data, no matter how big (and some data sets include >100 million people), can only provide us with estimates of association, albeit interesting and precisely estimated. These estimates of associations about the relative value of the available options fall short of meeting decision makers' needs to draw causal inferences. I do not trust instrumental variables and propensity scores and other fancy analytic tools designed to extract quasicausal inferences from these data sets. Their methodological sophistication cannot successfully overcome the limitations in the data set itself: big, yes, but riddled with error, incomplete (from key variables only available for some patients to massive silence about the biological, psychological, and socioeconomic context of each patient), and confounded. Although these hints can be useful, and sometimes very useful, too often Big Data is simply not Great Data.

Toward Big Science

Assessing the impact of sensible options on outcomes that matter requires very large randomized trials. These trials can only be assembled through massive participation. Here, it is not only the data that is big but also the scale of the collaboration across scientists and academic institutions, clinics and health systems, and patients and communities. This is not just Big Data. It is Big Science.

Big Science can help us characterize and estimate differences that matter not just with precision, but credibly. This goes beyond determining whether the options are different, to estimating the magnitude and nature of their differences across diverse patients, outcomes, and contexts. To answer these questions, Big Science must take place across geographies,

cultures, and models of care, a task that requires broad international collaboration in the conduct of large, multicenter randomized trials, prospective meta-analyses, and new designs we are yet to invent. The Table compares Big Science with other clinical care research approaches.

Just now, we are beginning to realize just how large Big Science has to be. For example, when the Patient-Centered Outcome Research Institute started funding trials, these planned to enroll only a few hundred patients. When Patient-Centered Outcome Research Institute pivoted to fund practical comparative effectiveness trials designed to meet the needs of decision makers,¹⁴ the planned size of these trials grew into the tens of thousands. To be feasible, work of this magnitude and complexity must take place within existing care settings, including practice networks dedicated to participate in a learning health system.¹⁵ To serve as a scaffold for the conduct of very large trials, Patient-Centered Outcome Research Institute invested in and assembled PCORNet (the National Patient-centered Clinical Research Network). The ADAPTABLE (Aspirin Dosing: A Patient-centric Trial Assessing Benefits and Long term Effectiveness), for example, was designed to make use of PCORNet to enroll 20 000 patients with coronary disease to compare 2 doses of aspirin,¹⁶ across outcomes important to patients and across pertinent subgroups. Big Science makes this scale of work possible.

From Competition to Collaboration

To get to Big Science, however, we must extirpate competition from clinical research. Private foundations, local research funding offices, and federal agencies all use competition to allocate resources and build their portfolios. Competition culls. Competition, mostly centered on obtaining research funding, produces very few winners and very many losers. Lack of funding is interpreted by academic institutions to mean that peers have judged your ideas as unworthy and funding your research as wasteful. I cannot find good evidence in support of the notion that this form of competition improves ideas, drives innovation, and fosters talent in science. Has research funded during periods of austerity yielded better science?

Consider the work necessary to produce a research proposal and that 8 or 9 out of every 10 of these proposals ends up unfunded and buried in the unreadable hard drive of an outdated laptop. Some researchers, like lottery players, play their odds by submitting and resubmitting as many applications, even mediocre ones, as they can. This produces the illusion of productivity and keeps funding agency staff and reviewers busy. With their creativity, time, and effort wasted, trained talented researchers writing proposals instead of conducting new experiments burn out and give up, leaving behind unanswered questions and unexplored ideas. Perhaps these investigators, their ideas, and approaches were the weakest. Perhaps our system also discarded the brilliant, the generous, and the pathbreaker.

Because for every winner there are many losers, the winner learns not to share their secrets, their contacts, their approach, and their resources. When researchers are in the same institution—a situation that should facilitate collaboration and Big Science—sharing is discouraged when the promotion of one may require the failure of competitors. Transparency

Table. Comparison of Big Science With Other Clinical Care Research Approaches

	Little Science	Big Data	Big Science
<i>Goal</i>	To meet needs of Food and Drug Administration regulators and to generate new knowledge	To support the needs of decision makers	To support the needs of decision makers
<i>Output</i>	Estimates efficacy on primary end point	Estimates associations, explore hypotheses, draw quasicausality inferences about comparative effects of alternative treatments, effects in subgroups, effects on infrequent or delayed outcomes	Estimate comparative effectiveness measured on outcomes that matter to patients, including burden of alternative courses of action
<i>Questions</i>	Address a goal important to the sponsor, regulator, or the investigator	Address what is possible given the available data	Address what is possible with generous collaboration and necessary for confident clinical decision making
<i>Participants</i>	Selected to optimize responsiveness and support favorable indications	Real-world patients receiving all or part of their care in participating health systems or care funded by participating payers and for whom data and follow-up is sufficiently complete	Patients receiving care in practice-research networks
<i>Comparisons</i>	Often placebo	Usual practice	Usual care with evidence-based alternatives that patients value
<i>Data and sources</i>	Per protocol	As-is in existing data sets	Per protocol
<i>Design</i>	Smallest fastest trials; multicenter with few participants per center; and meta-analyses of these trials	Use data collected from linked observational databases formed by the practice and administration of healthcare	Largest possible study (multicenter trials and prospective meta-analyses) to precisely estimate effects, beneficial, and harmful, across a broad range of patients and contexts
<i>Tactics to achieve efficiency</i>	Enroll responsive participants; use surrogate and composite end points; compare to placebo	Ask multiple questions against data collected in the course of providing or paying for care	Use data generated and collected in the course of providing (documented in records) or receiving care (clinician and patient collected/ reported)
<i>Outcomes</i>	Disease-oriented, surrogate markers assessed as soon as possible to obtain significant differences between arms	Surrogate markers, procedures, patient outcomes at sufficiently long follow-up periods by which data is sufficiently complete	Outcomes that decision makers (often patients) experience and value at sensible follow-up periods
<i>Size–funding relationship</i>	As big as funding affords	As big as available in datasets	As big as possible, requiring multiple funding sources
<i>Stakeholder involvement</i>	Optional, directed at making funding and recruiting feasible	Optional, directed at selecting analyses and disseminating results	Essential, from identifying question to disseminating results
<i>Scope of collaboration</i>	Across researchers with outreach to nonresearch clinicians and patient groups to optimize recruitment	Across researchers and data owners/ stewards	Across funders, healthcare practice networks, patient groups, researchers and dissemination agents, often internationally
<i>Confidence in the body of evidence</i>	Indirect evidence in terms of patients, comparators, and outcomes; limited precision, particularly outside of primary end point (pooling can improve but cannot overcome other biases); biased publication related to interest of sponsor and other conflicts	Direct and highly precise comparisons are likely, but hindered by residual confounding. Blurred lines between protocol-driven work and data-driven protocols. Publication and reporting bias highly likely	Direct and highly precise results, even across a range of patients and contexts, are likely. Publication bias is less likely thanks to public protocols and visible large-scale conduct in real world

and generosity, key scientific characteristics critical for Big Science, languish devalued and displaced by competition.

What It Takes to Do Big Science

The nature and magnitude of collaboration required for Big Science must follow from a fundamental change in the culture of research. Institutions will have to reward the collaborative

and generous scientist, one who excels at followership, fellowship, engagement, and inclusion. Big Science requires close partnerships within and between communities of research and clinical practice. The methods deployed must work at scale and cause minimal disruption in the process and experience of clinical care. They must balance rigor with adequate privacy protection. And, we will have to invent new ways of funding

to promote and support the work of collaborative multidisciplinary teams realizing best ideas and the practice networks that realize their plans.

Healthcare payers are exploring new care and payment models. We must consider the possibility that the biggest value proposition in healthcare involves caring for patients while we learn about how to improve care. Could payers reward care teams that collaborate with Big Science teams? It is possible that by breaking the budget silos of research and practice we can fund and sustain Big Science and improve evidence-based practice.

We need to be able to share protocols and ideas and need to be able to develop commons where that sharing happens freely and easily. The collaborative culture of Big Science should mitigate the fear that drives some researchers to claim ownership of ideas and data and facilitate collaboration in the secondary analyses of these complex and rich repositories, perhaps a better form of Big Data. Talented people sequestered in a room or thousands collaborating online could work toward the best possible research designs; communities of practice could pilot test and improve the feasibility of the protocol. The value proposition of Big Science, however, will be woefully incomplete unless it is able to fundamentally improve the lives of patients by translating evidence into practice. Some groups are leading the way through large scale collaborations that configure a learning system that is able to generate new evidence and improve quality of care.^{17,18}

Clinicians and patients will still need to struggle to carefully identify what is best for each patient at each junction. They will have to consider the available options, including the option to participate in clinical trials, until the best way becomes evident. I do not think that science can ever answer the question of what is best for each patient; that answer depends on the values, preferences, and dynamic context of each one. But Big Science offers a promising chance to ease the challenging practice of patient-centered evidence-based care.

This integrated learning system—one that produces and uses Big Science to advance patient-centered care—will need broad engagement of stakeholders. The time is right as it has become increasingly difficult to imagine research without engaging patients and caregivers and other stakeholders in all aspects of clinical research.¹⁹ A similar shift is taking place in practice where care is increasingly imagined as being cocreated with patients.²⁰ Leading healthcare institutions and individuals must give voice to the true magnitude of ignorance and uncertainty in which we seek to care and to improve care. The work of developing research-practice communities and their commons, of facilitating collaboration across ideas, protocols, work, and data, and of translating evidence into practice calls for inclusive and generous collaboration. The next big cultural challenge for biomedical science is to make generous collaboration fundamental.

Generous Collaboration

Science fairs sometimes inspire children to become scientists. What they learn in these fairs, unfortunately, is that a successful scientist is one that beats everyone else for recognition. Clinical researchers recognize the same rules apply

to grown-up medical science. Show up often with your not-so-innovative grantsmanship and compete to win. Keep your ideas, resources, and credit to yourself. These rules must change. To be successful, scientists focused on improving the care of patients must apply innovative craftsmanship. They must collaborate broadly while sharing generously and transparently without regard for credit. They must publish fully and liaise with others to ensure that science-based care reaches everyone who can benefit.

Big Science needs generous collaboration. The International Space Station, built and maintained by scientists from different countries orbits above, its fast and luminous path a monument to generous collaboration. In the bowels of Europe, the Large Hadron Collider is rewarding multinational financial and scientific collaboration with magnificent discoveries from the subatomic world. More insights into space and matter are being revealed thanks to the ingenuity, hard work, and generosity of collaborating people, institutions, and countries. The article describing the discovery of the Higgs boson listed the names of >5154 collaborators in its authorship byline in alphabetic order.²¹ Who discovered that particle? We all did. I hope the day will come when we can all celebrate the fruits of generous collaboration in medicine. A day in which Big Science helps patients and clinicians uncover what is best for the patient.

Acknowledgments

Dr Montori is grateful for the generous contribution to the ideas reflected here of his colleagues at the Knowledge and Encounter Research Unit at Mayo Clinic.

Disclosures

Dr Montori leads the Knowledge and Encounter Research (KER) Unit at Mayo Clinic; this research group has received over the last 12 years grant funding and contracts from nonprofit organizations for the conduct of systematic reviews and meta-analyses and for the formulation of practice guidelines and shared decision making tools based on these syntheses. Dr Montori and the KER Unit derive no other income from these activities. Dr Montori has no other financial relations to report.

References

1. May C, Montori VM and Mair FS. We need minimally disruptive medicine. *BMJ*. 2009;339:b2803. doi: 10.1136/bmj.b2803.
2. Hargraves I, Kunneman M, Brito JP, Montori VM. Caring with evidence based medicine. *BMJ*. 2016;353:i3530. doi: 10.1136/bmj.i3530.
3. Hargraves I, LeBlanc A, Shah ND, Montori VM. Shared decision making: the need for patient-clinician conversation, not just information. *Health Aff (Millwood)*. 2016;35:627–629. doi: 10.1377/hlthaff.2015.1354.
4. Krumholz HM. Biomarkers, risk factors, and risk: clarifying the controversy about surrogate end points and clinical outcomes. *Circ Cardiovasc Qual Outcomes*. 2015;8:457–459. doi: 10.1161/CIRCOUTCOMES.115.002245.
5. Ferreira-Gonzalez I, Busse JW, Heels-Ansdell D, Montori VM, Akl EA, Bryant DM, Alonso-Coello P, Alonso J, Worster A, Upadhye S, Jaeschke R, Schunemann HJ, Permanyer-Miralda G, Pacheco-Huergo V, Domingo-Salvany A, Wu P, Mills EJ, Guyatt GH. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ*. 2007;334:786. doi: 10.1136/bmj.39136.682083.AE.
6. May CR, Eton DT, Boehmer K, Gallacher K, Hunt K, MacDonald S, Mair FS, May CM, Montori VM, Richardson A, Rogers AE, Shippee N. Rethinking the patient: using Burden of Treatment Theory to understand the changing dynamics of illness. *BMC Health Serv Res*. 2014;14:281. doi: 10.1186/1472-6963-14-281.

7. Murad MH, Montori VM, Ioannidis JP, Jaeschke R, Devereaux PJ, Prasad K, Neumann I, Carrasco-Labra A, Agoritsas T, Hatala R, Meade MO, Weyer P, Cook DJ, Guyatt G. How to read a systematic review and meta-analysis and apply the results to patient care: users' guides to the medical literature. *JAMA*. 2014;312:171–9. doi: 10.1001/jama.2014.5559.
8. Murad MH, Montori VM. Synthesizing evidence: shifting the focus from individual studies to the body of evidence. *JAMA*. 2013;309:2217–2218. doi: 10.1001/jama.2013.5616.
9. Trinquart L, Abbe A, Ravaud P. Impact of reporting bias in network meta-analysis of antidepressant placebo-controlled trials. *PLoS One*. 2012;7:e35219. doi: 10.1371/journal.pone.0035219
10. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358:252–260. doi: 10.1056/NEJMsa065779.
11. LeBlanc A, Bodde AE, Branda ME, Yost KJ, Herrin J, Williams MD, Shah ND, Houten HV, Ruud KL, Pencille LJ, Montori VM. Translating comparative effectiveness of depression medications into practice by comparing the depression medication choice decision aid to usual care: study protocol for a randomized controlled trial. *Trials*. 2013;14:127. doi: 10.1186/1745-6215-14-127.
12. LeBlanc A, Herrin J, Williams MD, Inselman JW, Branda ME, Shah ND, Heim EM, Dick SR, Linzer M, Boehm DH, Dall-Winther KM, Matthews MR, Yost KJ, Shepel KK, Montori VM. Shared Decision Making for Antidepressants in Primary Care: A Cluster Randomized Trial. *JAMA Intern Med*. 2015;175:1761–1770. doi: 10.1001/jamainternmed.2015.5214.
13. National Academies of Sciences, Engineering, and Medicine. Refining the Concept of Scientific Inference When Working with Big Data: Proceedings of a Workshop-in Brief. Washington, DC: The National Academies Press, 2016.
14. Karanicolas PJ, Montori VM, Devereaux PJ, Schunemann H, Guyatt GH. A new "mechanistic-practical" framework for designing and interpreting randomized trials. *J Clin Epidemiol*. 2009;62:479–484. doi: 10.1016/j.jclinepi.2008.02.009.
15. Smith M, Saunders R, Stuckhardt L, McGinnis JM. *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. Washington, DC: National Academy of Sciences; 2013.
16. Johnston A, Jones WS, Hernandez AF. The ADAPTABLE Trial and aspirin dosing in secondary prevention for patients with coronary artery disease. *Curr Cardiol Rep*. 2016;18:81. doi: 10.1007/s11886-016-0749-2.
17. Forrest CB, Margolis P, Seid M, Colletti RB. PEDSnet: how a prototype pediatric learning health system is being expanded into a national network. *Health Aff (Millwood)*. 2014;33:1171–1177. doi: 10.1377/hlthaff.2014.0127.
18. Crandall W, Kappelman MD, Colletti RB, Leibowitz I, Grunow JE, Ali S, Baron HI, Berman JH, Boyle B, Cohen S, del Rosario F, Denson LA, Duffy L, Integlia MJ, Kim SC, Milov D, Patel AS, Schoen BT, Walkiewicz D, Margolis P. ImproveCareNow: The development of a pediatric inflammatory bowel disease improvement network. *Inflamm Bowel Dis*. 2011;17:450–457. doi: 10.1002/ibd.21394.
19. Selby JV and Slutsky JR. Primary care research in the patient-Centered Outcomes Research Institute's Portfolio. *Acad Med*. 2016;91:453–454. doi: 10.1097/ACM.0000000000001116.
20. Batalden M, Batalden P, Margolis P, Seid M, Armstrong G, Opari-Arrigan L, Hartung H. Coproduction of healthcare service. *BMJ Qual Saf*. 2016;25:509–517. doi: 10.1136/bmjqs-2015-004315.
21. Aad G, Abbott B, Abdallah J, Abdinov O, Aben R, Abolins M, et al; ATLAS Collaboration; CMS Collaboration. Combined measurement of the Higgs Boson Mass in pp collisions at sqrt[s]=7 and 8 TeV with the ATLAS and CMS experiments. *Phys Rev Lett*. 2015;114:191803. doi: 10.1103/PhysRevLett.114.191803.

KEY WORDS: biomarkers ■ caregivers ■ coronary disease ■ creativity ■ evidence-based medicine

Big Science: Research Collaboration for Evidence-Based Care Victor M. Montori

Circ Cardiovasc Qual Outcomes. 2016;9:688-692; originally published online November 8, 2016;

doi: 10.1161/CIRCOUTCOMES.116.003358

Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Copyright © 2016 American Heart Association, Inc. All rights reserved.

Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the World Wide Web at:

<http://circoutcomes.ahajournals.org/content/9/6/688>

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Circulation: Cardiovascular Quality and Outcomes* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the [Permissions and Rights Question and Answer](#) document.

Reprints: Information about reprints can be found online at:
<http://www.lww.com/reprints>

Subscriptions: Information about subscribing to *Circulation: Cardiovascular Quality and Outcomes* is online at:
<http://circoutcomes.ahajournals.org//subscriptions/>