Here's One Way to Do Better Science

written by Robert McNutt, M.D. | May 30, 2019

Clinical research with randomized trials (RTs), as opposed to basic or bench research, is the science of *comparison*. RTs ask a fundamental question: Is "x" better than "y"? They do more than observe how treatments work; they also require methods that control the research environment. Finding an independent contribution of one action over another demands *random, stratified populations in order to find truthful differences*, as medical care advances on these differences. But the way that patients are typically recruited for RTs can undercut the validity of the study's findings. I propose we take a different approach, one I call "Gallup Research Medicine."

Current Study Recruitment Is Far From Systematic

First, let's look at how patients are often recruited for study. Physicians, recruited by researchers, ask patients to participate in trials. Having watched many physicians recruit, I find they ask patients they think will participate, excluding others for unsure reasons. They rarely recruit in a systematic manner (randomly or consecutive sampling). Sometimes physicians are busy and recruit patients only when they are not inundated. Sometimes they abdicate the responsibility of recruiting to others in their office. Physicians might even be paid to include

patients. I know of one physician, offered more than $5,000 per patient to enter two patients in a trial of a device in the intensive care unit (ICU), who daily scoured the ICU looking for the perfect patients to include, while bypassing eligible patients.

For all these reasons, physicians should not choose who is in a study; they have other things to do. Sometimes, researchers recruit patients by getting lists of patients from a physician and, each day, contacting those coming for a visit. But visits are not random events, and recruiters are not always available when patients are. So, even bypassing physician recruitment is not good enough; a better way to recruit is needed.

Gallup Polls Offer Insight into a Better Recruitment Strategy

Here is an alternative approach for recruiting and studying more representative samples of patients in RTs. A little background: In the 1990s, I attended a lecture showing a new idea for an electronic health record at Harvard. Their system at the time included about 90 sites of care. Each time a person accessed care at a site, they were flagged as having a visit, and the visit (event) was stored in an event database. The data in this event database only included person, place, time and a "pointer" from the event database to the result database at that site. For example, if I went to site X for a test, the test result was housed in X's database, but my visit event was sent to a separate database.

Were I to see a physician, then, my event database would be queried and data from disparate sites loaded to a health record interface. Had I visited ten sites, the data from those would populate my record. When I left the visit, the interface would disappear, my visit would head to the event database, and, separately, the results of that event would head for the site database.

Now, let's turn to Gallup polls. Gallup's methodology is based on a similar conceptual idea to the electronic record I saw decades ago. They have an "event database" and use that event data to point to and gather further data from each of those events. The events are telephone numbers. The "stored data" on the other end of the phone number is a person. On a daily basis, 350 days per year, Gallup calls, *randomly*, 500 phones and asks for information.

These two experiences—(1) a virtual electronic health record informed by data housed at different sources but collated by an event database, and (2) poll data from a random sample of an entire population but collated by an event database—are essentially the same idea and suggest a way to produce better RTs. If my event and site-specific data can be collated, and if Gallup can randomize from full populations, then all women with breast cancer or heart disease, all men with prostate cancer, or any person with any defined illness can, likewise, be

enrolled in disease-specific event databases. Should a study be needed, a random sample of diseased people could be queried.

How "Gallup Research Medicine" Could Work

Here's an example. Suppose I am uncertain if the woman discussed in a previous blog with triple positive (ER/PR/HER2) breast cancer should get the expensive, dangerous anti-HER2 drug. To answer this question, I would query a breast cancer event database at every health care site, produce a *nationwide breast cancer event registry*, and then call a random sample of women in this registry and ask for permission to gather their data for a research study.

If a woman agreed, we could then collate *disease specific data*, run an exclusion/inclusion algorithm to limit the population to triple positive women, for example, thereby creating a random, full population sample for study. At each site, remember, the event database points to data such as receptor status, stage, genetic markers and other data measured on all women. A research team would determine what data is needed.

The barriers for this idea are not technical. While I am not a computer wiz, I might start with an event database of person, place, time and disease condition. Alternatively, the event database could just be a disease, a pointer to the person, places, times of their care, and another pointer to the site data.

What about calling a random sample? Gallup calls 500 people daily, 350 days per year. This means they contact 175,000 people a year. I am unsure how many people make calls, but if this single company can contact so many, imagine how many people could be contacted with a coordinated effort from a government research agency, or multiple polling companies. To give some research context, 242,476 women developed breast cancer in 2015, (last year analyzed). Of these, about 25 percent are triple positive, or about 60,000. It would be a straightforward task to contact this number of people, either in full or as a random sample.

Government Has the Research Infrastructure to Lead Such an Effort

While this may not be difficult to organize, I suggest government should lead. Government has the infrastructure for research with rules for human subject research. They can encourage each hospital or electronic record company to keep disease-specific event databases. In fact, our government is presently incentivizing the sharing of data; this research effort could be a pilot project.

The leadership for research planning might be best organized at a regional level, however. A present research organization model with 60 sites in the US is Clinical and Translational Research Centers (CTRC). These centers have recruitment offices, study design expertise, and management teams that could keep the collated disease registries, perhaps educate patients and providers, present updates on outcomes via transparent measures, and revise if needed as studies progress. In addition, Institutional Review Board considerations, informing patients of research intent, and ethical considerations could be managed at the CTRC sites, as they now have those activities up and running.

These ideas are not the only potential ways to advance larger sample, appropriately stratified research studies. The idea behind an event database with pointers to on-site data is to limit problems with data conversions between disparate systems of care. Two systems would not have to share the same data structures, needing only to send targeted research data from their sites using their own data teams.

However, the first production version of Fast Healthcare Interoperability Resources (HL7 FHIR®) just launched in 2018; HL7 FHIR standardizes the structure of data for sharing. Already, disparate sites of care are developing uniform ways to transfer data using this software. Apps for HL7 FHIR could, without event databases, run queries of data at all electronic record sites to build disease registries. I am sure there are other ideas, but every effort should be made to advance the science of the RT.

Imagine a future of research using random samples of full populations for study, rather than haphazard samples and incomplete groups of prognostically different patients that are now the norm. While technical considerations are solvable, there are likely legal or political considerations, and current business models of conducting RTs may be upended. However, doing high quality research should be such an important public health priority that we and our government should demand it.

Founded as ICLOPS in 2002, Roji Health Intelligence guides health care systems, providers and patients on the path to better health through Solutions that help providers improve their value and succeed in Risk. Roji Health Intelligence is a CMS Qualified Clinical Data Registry.

Image: Jacek Dylag