By Ajai Raj

The current state of clinical

(http://www.clinicaloncology.com/aimages/2016/OncSE0616_0003_a_425.jpg)

oncology gives many experts cause for optimism. Recent developments, including personalized medications and immunotherapy, show promise as the next generation of treatment. Before these new therapies can be given to patients, however, they must run the gauntlet of multiple clinical trials to verify their efficacy and safety. And, just as treatments are increasing in complexity, experts say the way trials are conducted must also adapt to meet the challenges of testing these approaches effectively.

From inefficient trial design to a widespread lack of statistical sophistication among researchers and even misconduct (see page 5), the outlook for clinical research has never been more troubled, according to experts in the field. The hoarding of data by individual research teams from the cold light of widespread scrutiny threatens a pillar of scientific research in general and oncology in particular: reproducibility.

But if reproducibility is the disease, experts agreed, data sharing is the cure. In fact, unrestricted sharing of data and the code used to reach scientific conclusions has the potential to cure many ills,

Therapy-Promising-for-Multiple-Myeloma/39011)

Less-Intensive Regimen Posed for Some BL Patients (/Hematologic-Malignancies/Articl 16/Less-Intensive-Regimen-Posed-for-Some-BL-Patients/38856)

Tweets by @ClinOncNews

ClinicalOncologyNews
@ClinOncNews

Thyroid: Survival Rates,MDM, and More
bit.ly/2idXhF0

11 Jan

ClinicalOncologyNews
@ClinOncNews

#ASH2016 —Ruben Mesa, MD Discusses

Embed   View on Twitter

Most Read

Early Detection of Breast Cancer-Related Lymphedema (/Breast-Cancer/Article/12-16/Early-Detection-of-Breast-Cancer-Related-Lymphedema/38864)

BREAST CANCER (/Section/Breast-Cancer/553)

including the lack of statistical knowledge among oncology researchers and a host of other "stupid mistakes" that run rampant in the field, experts said.

New Treatments, New Trials

"A lot of clinical trials don't tell us everything we need to know until we wind up seeing the effects of drugs in much, much larger populations over a long time period," such as in Phase IV trials, said Andrew von Eschenbach, MD, a former director of the National Cancer Institute (NCI) and a former FDA commissioner, who is now a senior fellow at the Milken Institute, a Santa Monica, Calif.-based think tank.

(http://www.clinicaloncology.com/aiimages/2016/OhcSE0616_00

Andrew von Eschenbach, MD

Dr. von Eschenbach said tools are available now that could alter the traditional paradigm of proceeding through Phase I, II and III trials—to establish toxicity, efficacy and efficacy relative to existing treatments, respectively— before the FDA grants its approval to a new agent.

"When the original paradigm was put into place, we had to work with physical manifestations of disease, but those didn't do anything as far as telling us what to do about the problem," he said. "It was very much empiric, trial and error."

The level of understanding has evolved with the available technology, to the point that researchers can begin to develop new frameworks for thinking about trials, one that would meet current regulatory requirements and provide the confidence needed to make decisions, he noted.

"For example," Dr. von Eschenbach said, "we know that many of these drugs now are no longer disconnected from diagnosis, and that in the diagnostic part of the equation—the biomarkers, our

New CAR T-Cell Therapy Promising for Multiple Myeloma (/Hematologic-Malignancies/Article/01-17/New-CAR-T-Cell-Therapy-Promising-for-Multiple-Myeloma/39011)

HEMATOLOGIC MALIGNANCIES (/Section/Hematologic-Malignancies/149)

COA Report: Practices Closing, Cost of Cancer Care Increasing (/Current-Practice/Article/12-16/COA-Report-Practices-Closing-Cost-of-Cancer-Care-Increasing/38841)

CURRENT PRACTICE (/Section/Current-Practice/155)

SABCS 2016: No DFS Benefit From Extended Adjuvant Endocrine Rx With Letrozole in HR+ Breast Ca (/Breast-Cancer/Article/12-16/SABCS-2016-No-DFS-Benefit-From-Extended-Adjuvant-Endocrine-Rx-With-Letrozole-in-HR-Breast-Ca/38804)

BREAST CANCER (/Section/Breast-Cancer/553)

Vogl, NY (/Section/Vogl-NY/556)

Myeloma–Time To Push for a Cure With MRD End Points (/Vogl-NY/Article/11-16/Myeloma-Time-To-Push-for-a-Cure-With-MRD-End-Points/38453)

It's time for a push toward cure of MM. The best available landmark for such a push...

understanding of the genetic abnormalities that are associated with tumors—we're beginning to see that there are subsets of patients that we used to lump all together as 'breast cancer' or 'lung cancer,' that we can now start separating into more discrete and more defined groups," such as tumors with HER2 and ALK mutations. So suddenly, we begin to see that we may be able to control for certain variables that enrich your population, and therefore randomization becomes less of a need because we can really now manage the critical or key variables in what was a heterogeneous population."

ADVERTISEMENT

During his tenure at both the NCI and the FDA, Dr. von Eschenbach helped to create the concept of a Phase 0 trial, in which a small amount of a drug is administered to a few patients. The purpose is not to test toxicity or efficacy, he said, "but simply to prove that this drug, designed to hit a certain target, is in fact hitting that target, that mechanism of action that is associated with the pathology of the disease state. You could then assume that there's going to be clinical efficacy, clinical benefit subsequent to that."

Rethinking Significance and Banishing Bias

As the field moves forward, some of the more basic assumptions and metrics on which the authority of clinical trials rests are being called into question. For instance, although large sample sizes generally are desirable, increasing the confidence with which study conclusions can be applied to wider populations, large studies also have a potential for abuse, said Harold Sox, MD, a clinical epidemiologist and director of research portfolio development at the Patient-Centered Outcomes Research Institute (PCORI), an independent nonprofit located in Washington, D.C. PCORI's mission is to improve the quality and relevance of evidence available to help patients, caregivers, clinicians and other stakeholders make informed health decisions.

Extended Adjuvant AIs: Safe and Active, but Not Much Good (/Vogl-NY/Article/09-16/Extended-Adjuvant-AIs-Safe-and-Active-but-Not-Much-Good/37892)

Paul E. Goss, MD, PhD, presented the 6.3-year analysis of the MA.17R breast cancer trial at the...

Archives (/Archive)

"If you compare two drugs in a very large study, if there's only a small difference between them, and if the study is big enough, you'll detect the difference well enough to know that it's real," Dr. Sox said. "But because the study is so large, you could get yourself into a situation where you've demonstrated statistical significance, but the difference is not big enough clinically for patients to appreciate."

As an analogy, Dr. Sox cited a meta-analysis of four studies comparing surgery with nonsurgery for chronic nonspecific lower back pain. The study found that the average disability rating, measured on a scale of zero to 100, was four points better for surgery (Spine 2007;32:816-823).

"Most people think that the patients can't really tell if they're better when the difference is that small," Dr. Sox said. "Yet in this particular study, that difference of four points on a scale of zero to 100 was statistically significant, because the authors of the meta-analysis pooled the results of four studies and got a pretty good sample size." Furthermore, using just the average improvement in disability as the measure of effectiveness may disguise a large reduction in disability for a small number of patients. A focus of current research is how to identify patients who stand to gain a lot from surgery.

"The sample size was very large, way up in the thousands, and so the difference between the two drugs was statistically significant, but it was a tiny difference," he said.

Many researchers, both within and outside oncology, now agree that although statistical significance can help boost confidence in trial findings, its utility is limited, particularly when appraised in terms of P value.

"The problem is that demonstrating statistical significance doesn't mean that we now know that the drug works," said Kay Dickersin, PhD, the director of the Center for Clinical Trials and

Evidence Synthesis at Johns Hopkins University, in Baltimore. "It just tells us the probability that the results are due to chance."

Calling the P value "convoluted and difficult to understand," Dr. Dickersin said that, as a rule, epidemiologists prefer to use the confidence interval instead. "Say you have a 99% confidence interval—that gives you the range in which 99% of the time, you'd expect the true value to fall," she said. "It's a much more informative way of testing the reliability of an estimate."
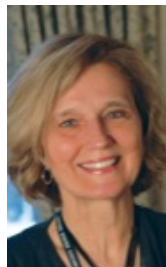
(http://www.clinicaloncology.com/aimages/2016/OncSE0616_000

Kay Dickersin, PhD

Indeed, some journals are moving away from having authors report P values in manuscripts, and one, Basic and Applied Social Psychology, last year banned the statistical measure entirely on the ground that it "fails to provide the probability of the null hypothesis, which is needed to provide a strong case for rejecting it."

Another problem with the P value is the extent to which some researchers go to reach the hallowed less than 0.05 threshold that presumably indicates significance. Such "P hackers" effectively put the statistical cart before the horse by dredging their data for evidence for an association without first coming up with a reason for the association itself. As one author put it: "The problem is that you introduce bias when you choose to collect more data (or analyze the data differently) only when the P value is greater than 0.05. If the P value was less than 0.05 in the first analyses, it might be larger than 0.05 after collecting more data or using an alternative analysis. But you would never see this if you only collected more data or tried different data analysis strategies when the first P value was greater than 0.05" (Naunyn Schmiedebergs Arch Pharmacol 2014;387:1017-1023).

Echoing Dr. Sox's comments about small differences, Dr. Dickersin said that while statistical significance does not always indicate clinical significance, measuring clinical significance on its own presents serious hurdles. "That's a tough one—after all, what is clinical significance?" she said. "What's significant for one person might not be for another. It's a question of tradeoffs, of personal values."

Another challenge in oncology trials is the erosion of confidence caused by bias. "There's no question that there are reporting biases in oncology, as in any field of study where we have looked," Dr. Dickersin said. These biases come in many forms, she noted, including publication bias, in which favorable results are published and unfavorable ones held back, and selective outcome reporting. A common example of the latter situation is when the primary outcome measure of a trial does not reach statistical significance but a secondary outcome measure does—prompting the investigators to reframe their study to focus on the secondary outcome and ignore the original measure.

Dr. Dickersin, who has participated in studies examining publication bias and selective outcome reporting, said the problem is more widespread than has originally been thought. "People have been extremely worried about drug companies modifying results to make drugs look good," she said. "But we're now seeing that academics have problems, too."

The Case for Sharing Data

The single most potentially effective corrective for these and other biases, said Dr. Dickersin and others, is to register all trials and make open access to clinical trial data the norm. Doing so would allow for more meta-analyses, which provide more confidence in findings than individual studies by themselves.

"We want to know what the truth is regarding how well a treatment works, so we really need to know what's happened in every trial," Dr. Dickersin said.

"Patients have given their consent to participate, believing they're contributing to knowledge. So scientists should be sharing all the knowledge they have obtained with others, including the public.

"There's no reason we should keep knowledge to ourselves," she emphasized. "That's not what the patient thinks we're doing. They're participating in trials to contribute to knowledge, and they don't want us to hide any part of what we know."

Indeed, although Dr. Dickersin's position seems incontrovertible in light of medical ethics, powerful voices have spoken out against the open sharing of data. For instance, this January, theNew England Journal of Medicine published an editorial by editor-in-chief Jeffrey Drazen, MD, and Dan Longo, MD, arguing that data sharing posed the danger of breeding "research parasites," a hypothetically harmful class of researchers who Drs. Drazen and Longo characterized as "people who had nothing to do with the design and execution of the study but use another group's data for their own ends, possibly stealing from the research productivity planned by the data gatherers, or even use the data to try to disprove what the original investigators had posited" (N Engl J Med 2016;374:276-277). The editorial was roundly pilloried by scientists on Twitter, who pointed out attempting to poke holes in hypotheses is at the heart of science. The backlash was so intense that Dr. Drazen was moved to write a second editorial in which he substantially tempered his views on data sharing (N Engl J Med 2016 Jan 25. [Epub ahead of print]).

In addition to being an ethical imperative, open access to clinical trial data is simply common sense, the lack of which can have disastrous—and easily avoided—consequences for the trustworthiness of studies, said Keith Baggerly, PhD, a professor in the Department of Bioinformatics and Computational Biology at the University of Texas MD Anderson Cancer Center, in Houston.

"The question I have been wrestling with for the last few years is, how can we improve the reproducibility rates of preclinical studies—primarily in genomics, the area I spend most of my time in," Dr. Baggerly said. "And a big chunk of that has been focused on the question of saying, 'Congrats, we've reached a new threshold.' The data sets we're working with are so large, and the analysis sets are so large, that we can't hold all of what we're doing in our brains. Because of this, we may not see problems as they emerge, may not be able to see things and say intuitively, 'Oh, that can't be right.'

"What that means is that written descriptions and discussion sections are likely to be inadequate to help others reproduce the research," he explained. "We've reached the point where we need to be able to make available to others both the raw data and the code scripts used to produce the results."

The scale of the problem of reproducibility in oncology—or the lack thereof—was made clear in a 2012 article inNature (2012;483:531-533) by a group from Amgen. Researchers in the hematology and oncology department at the company sought to replicate 53 "landmark papers" related to a certain line of research to explore the commercial potential of the discoveries. Only six were found to be replicable.

"That's not good enough," Dr. Baggerly said. "We should be setting the bar higher, expecting more" from preclinical as well as clinical studies.

Similarly, a 2013 study inPLoS One found that nearly 60% of researchers at MD Anderson who responded to a survey reported having been unable to replicate published data on at least one occasion (2013;8[5]:e63221). According to the authors of that report: "Overall, only 33.3% of respondents were ever able to explain or resolve their discrepant findings."

The failure to generate reproducible results may be a substantial tax on science. According to one recent estimate, the nation wastes roughly $28 billion each year on unreproducible research—a number that, if accurate, amounts to basically an entire NIH budget annually in squandered resources.

Dr. Baggerly said that the collection, archiving and open access of information provided by ClinicalTrials.gov is a model that ought to become more widespread. A large part of the model's effectiveness comes from incentivizing the sharing of data. Dr. Baggerly thinks a similar incentive could be implemented fairly easily on the publication side.

Studies of the role of genomics in oncology are fraught with "lots of potential statistical pitfalls," said Lisa McShane, PhD, the chief of the Biostatistics Branch of the Biometric Research Program, part of the Division of Cancer Diagnosis and Treatment at NCI. Dr. McShane has dedicated a significant portion of her career to developing and improving guidelines for data sharing, particularly for studies using so-called "omics" data, gene expression microarray data, proteomics data and other examples of what she calls "higher-dimensional biomarker measurements."

"We'd had some difficulties, particularly with genomic predictors, or mathematical models or functions that are built using omics data," she said. "When you start with a lot of variables and build a complex function, it's easy to do what's called 'overfitting' the data, which means you effectively built your function to pick up noise in the data, instead of true signal. If you don't know the proper procedures, it's easy to come up with a predictor that will work well on your own data set, but won't work anywhere else."

Too often, Dr. McShane added, she and her colleagues found flawed genomic predictors being used in clinical trials, due to errant statistical methodology. In some cases, she said, researchers did not know how to handle large data sets, which