The answer, of course, is learning how to share. In medicine, the question is when, and how, to share data. The hope (and hype) of Big Data is that massive datasets can help researchers and companies slice and dice cancer into smaller, more manageable patient populations where companies can rapidly develop, test, and launch new cancer targeted drugs and diagnostics. The goal, of course, is to generate much better patient outcomes, at much lower cost, than the status quo.

There’s also an increasing demand for companies to develop better metrics for demonstrating their products’ value to patients and the overall health care system. Sharing data (with appropriate protections for intellectual property) can help companies, researchers, and
oncologists refine their understanding of cancer’s many molecular pathways—and fine-tune the algorithms needed identify key vulnerabilities, while matching drugs to the patients most likely to respond (essential to improving the “value proposition” for patients, payers, and innovators).

It’s hard to overstate how devious and unpredictable cancer can be. Cancer cells have much in common with viruses and bacteria, at least in terms of their ability to evolve to adapt to drug treatment. Like HIV/AIDS, successful cancer therapies that cure or keep the disease in check indefinitely are likely to require precise cocktail therapies that overwhelm cancer’s ability to mutate and survive.

(commons.wikimedia.org/wiki/File:Wikipedia_-_Cancer_Survivor.jpg)
Researchers know this. The problem is that no single institution, or company, has all the pieces required to build the “knowledge network (https://www.ucsf.edu/news/2011/11/11004/nas-report-calls-building-biomedical-knowledge-network-drive-precision-medicine)” needed to defeat cancer in real time. Winning the war on cancer will take unprecedented collaboration: across hospitals, drug companies, regulators and insurers, all of whom possess key parts of the data puzzle.

This is where projects like ASCO’s CancerLinq (http://www.asco.org/quality-guidelines/cancerlinq), NIH’s Biomarker Consortium (http://www.biomarkersconsortium.org/index.php), and Project Data Sphere (https://www.projectdatasphere.org/projectdatasphere/html/home.html) come in. Project Data Sphere is the most recent, and in some ways, most interesting of the collaborations because it’s leveraging clinical trial data from
some of the industry’s leading companies—normally fierce competitors—to build a research platform that will be, for most intents and purposes, open to all comers (anyone with a plausible research question which might lead to a better understanding of cancer).

**A Golden Era Of Cancer Drug Development—But Too Many Opportunities and Lives Are Still Lost**

Wait, you might ask, aren’t we already in the midst of a golden era of cancer drug development? It certainly looks like it.

Every week seems to bring announcements of promising new oncology targets, better drugs in the pipeline, and better patient outcomes (including less toxicity). The FDA’s Oncology Division, inside the Center for Drug Evaluation and Research (CDER), has a good reputation among sponsors as being more flexible and innovative than some other divisions. Pricing for oncology products is at a premium, and new drugs routinely command six figures—attracting yet more investment and research.

All this is reflected in drug companies’ pipelines. PhRMA estimates that out of the more than 900 biologic medicines in development in 2013, a staggering 37 percent (the largest category by far) were for cancer or related conditions. And about 30% of all drugs in development and projects in 2011 were focused on oncology. From 2005-2012, close to 20% of all FDA approved drugs were for oncology indications. Today, cancer patients have greater hope than ever of living longer, better lives after a cancer diagnosis—and even, in a few cases, finding an outright cure.

But don’t crack the champagne quite yet. Oncology drug development remains plagued by lower success rates than other indications (although that seems to be changing, thanks to biomarker targeted trials). Meanwhile, those vaunted new technologies (genomics, proteomics, etc.) being brought to bear on oncology come with their own added costs. Only a sliver (3-5%), moreover, of all U.S. patients participate in clinical trials, with far too many oncology trials (about 20%) failing to accrue enough patients to reach completion—for reasons that have nothing to do with efficacy or adverse events. That’s a lot of valuable time and resources wasted.
In short, oncology drug development is, in many ways, getting more expensive and challenging, even as the underlying technology gets better.

Rising co-pays associated with new, high priced cancer meds are also forcing more patients to avoid effective therapies. As a result, more hospitals and physicians are pushing back against cancer drug pricing, in some instances even calling for outright price controls.

Like the rest of the health-care system, companies developing new cancer treatments are going to have to deliver better outcomes for every dollar spent, or face increasing calls for government to step in and set prices by fiat. Something has to give.

**Project Data Sphere**

In the not too distant past, companies were confident—perhaps too confident—that they could push an oncology drug through to the clinic with their own analytics and research platforms. That confidence is waning somewhat, especially as companies and researchers come to grips with cancer’s “high dimensionality” problem.

There may be a relatively small number of master regulator genes common to many cancers, but hundreds of other aberrant mechanisms exist—driving tumor proliferation and growth—which are only found in small subsets of patients (a few percent here and there), that can, in turn, also mutate in response to drug treatment. (Think of this as evolution in real time, as drugs select for the survival of the most drug-resistant tumor genotypes.)
In practice (http://www.nature.com/nrclinonc/journal/v11/n2/full/nrclinonc.2013.244.html), this means that we have too many drugs and cancer targets to ever run them all through traditional clinical trials, let alone test them in combination treatments. Nor do we have enough money or patients to even attempt it.

That’s where data sharing initiatives like Project Data Sphere (PDS) can make a difference.

An initiative of the CEO Roundtable on Cancer’s Life Sciences Consortium, PDS has created a repository for sharing and analyzing (de-identified) patient level data from Phase III comparator trials run by academic medical centers and drug companies.

What makes this project so remarkable is that the CEO Roundtable managed to get companies who are notoriously protective of their IP to share data. PDS launched with datasets provided by AstraZeneca, Bayer, Celgene, Janssen Research (an affiliate of Johnson and Johnson), Memorial Sloan Kettering, Pfizer, and Sanofi—and is looking to add more in short order. The SAS Institute (http://www.sas.com/en_us/home.html/) is also donating analytic tools (including allowing natural language queries) to registered users of the Web site. Access to the datasets is free too. As the editors at FierceBiotech (http://www.fiercebigdata.com/story/project-data-sphere-leverages-cancer-research-collaboration/2014-04-14) put it recently, everyone from a “high-school whiz kid” to the stats geeks at Google can turn their minds to chopping and dicing the PDS data sets.

Opening up industry comparator arms in one place could have some big payoffs. The first would be to improve clinical trial design, by identifying patients who either are high responders to current comparator treatments (and figuring out why), or identifying patient subtypes who fail rapidly on established therapies—and thus represent significant pockets of unmet medical need. Enrolling those patients in new trials would help increase success rates and represent real improvements in the status quo, thereby ensuring that we’re not giving new drugs to patients who won’t benefit from them.

Pooling enough historical data on disease progression and comparator treatment response could eventually lead to even bigger changes in clinical trial design—like dispensing with comparator arms altogether in favor of historical controls (although this will require a lot of buy-in from FDA regulators).
Larger pooled datasets could also be fed into machine learning platforms to create better protocols for the best treatment/patient matches, suggest better combination therapies to ward off drug resistance, and identify promising biomarkers. Virtual clinical trials run through these types of platforms could, likewise, help increase trial success rates and shave years off of drug testing.


Which is the “right model”? Maybe all of them. Or maybe the right one hasn’t been invented yet.

Over time, researchers will figure out what’s working and what’s not—and groups will naturally figure out where their comparative advantage is, or merge with other groups to better leverage scarce resources.

Today, the Big Data revolution in medicine is still more talk than substance. Interestingly, PDS may change that by launching a series of mini-“challenges (http://www.nccancerhospital.org/news/ceo-roundtable-on-cancer-project-data-sphere)” for its data sets (like identifying molecular signatures for particularly aggressive prostate cancers). Think of it as a series of fast, cheap “X-Prize” style-competitions.

It’s this focus on developing testable metrics through a crowd-sourced approach that could be PDS most valuable contribution to the shared learning space in oncology.

Advocates of mandatory data sharing across Big Pharma frame it as an exercise in transparency. But forcing companies to broadly disclose clinical trial data for innovative products risks upending current research incentives, which hinge on strong intellectual property rights. As my colleague Peter Huber points out in a recent issue of City Journal (http://www.city-journal.org/2013/23_4_genetic-data.html):
The successful development of a pioneering drug reveals key information about the molecular mechanics of a disease and a good strategy for controlling it. Armed with that knowledge, competitors can then modify the drug’s chemistry just enough to dodge the pioneer’s patent and rush in with slightly different drugs developed at much lower cost. In the end, the pioneer can easily be the only player that fails to profit from its own pathbreaking work.

In other words, if we free the wrong data without the right IP regime, we could find ourselves with a high-tech version of the tragedy of the commons.

Putting the focus on actionable data to improve cancer care and cancer drug testing through comparator arm data can build the trust required for more ambitious programs—while policymakers sort out some of the more complex IP issues involved.

Every player in the cancer space is going to have to learn how to collaborate to improve cancer outcomes at lower cost. Data sharing is one important part of that challenge.

At the end of the day, it’ll cost a lot less in lives and dollars to beat cancer with the right data.

INVESTOR’S NOTE: The oncology space remains a growth area with strong pipeline prospects, premium pricing, and increasing demand from an aging global population. Companies that prosper in this space will combine a commitment to oncology with strong academic partnerships to stay on top of advances in basic science – collaborations like Sanofi and UCSF (NYSE:SNY), Pfizer (NYSE:PFE), Eli Lilly (NYSE:LLY), and Merck (NYSE:MRK) have also created a nonprofit collaboration called the Asian Cancer Research Group, which is working with the China-based genomics group BGI, the University of Hong Kong, National University of Singapore and others to develop a better understanding of cancers in Asian populations. Other oncology winners will include companies with strong companion diagnostic platforms like Roche (ROG:VX). Rising costs and new bundled payment pathways for oncology will also increase payor demand for platforms like IBM’s
(NYSE: IBM) Watson oncology project, developed in collaboration with Memorial Sloan Kettering and Wellpoint (NYSE: WLP).