

## University Senate Research Committee Meeting

(online)

17 February 2023

1:00 PM

**In attendance:** R Rutenbar, K Wood, S Sant, Z Xia, P Morel, M Holland, B Yates, D Reed, U Ndoh, J Silverstein, B Shieu, M Arlet, C Hay, R Kear, M Kenney

**Approval of Minutes:** The Jan 20, 2023 meeting minutes were approved without changes.

### Old items of discussion:

S Sant: The breakdown of faculty in the different levels of appointment and tenure track streams, and numbers of those with the “research” prefix, are not currently available from the Department of Pharmacy due to ongoing negotiations with the labor union.

### Update on Pitt’s external sponsored research rankings

Rob Rutenbar, Senior Vice Chancellor for Research  
(presentation)

Different metrics are used for school rankings, each measuring a different thing across a different timeframe.

NIH funding to institutions. The NIH publishes each January a listing of all its funding to schools and enterprises. NIH produces this list as it works to close out its books from the previous calendar year, explicitly stating that the data is preliminary and not ranked. But everyone else uses the numbers for ranking. Based on this NIH list, the Pitt institution ranks number 3 among universities in the NIH funding it receives. Pitt has always been in the top 10, and in recent years in the top 5, but this is the first time Pitt is number 3, which is 6 places higher than before. The listing also showed that NIH funding increased almost \$80 million year over year. This is a remarkable accomplishment for Pitt. Only Hopkins and the University of California at San Francisco (UCSF) are ranked higher than Pitt. This also says something because Hopkins is enormous and UCSF is almost exclusively health sciences, which largely explains their superior rankings.

NIH funding to medical schools. Another listing of yearly NIH funding to medical schools (not the entire institution) has Pitt achieving the number 6 ranking, which is also its best ranking in many years. There are places that own their own medical school: Hopkins, Michigan, San Diego. Then there are places where the university, not the affiliated hospital, owns the medical school: Pitt and Vanderbilt. There is also a complicated echo system of hospitals, where architecture of the system causes the money to go to the hospital. Harvard is one of these and, hence, not on the NIH list. If its funding went directly to its school, which it doesn’t, then Harvard would be at the top of the NIH list.

R Rutenbar's team did a data dive and found a \$77 million funding increase for Pitt. When they looked at NIH awards to Pitt, grouped by \$\$ range (<1 million, 1-5 million, 5-10 million, 10+ million) and found that in the \$10+ million range most of the funding decrease was for Covid-related work. In the \$5-10 million range, our numbers showed growth in the very large awards, which result from teaming and collaborations across school boundaries (health sciences and lower campus), which is where the NIH budget has grown to accommodate large, collaborative, grand challenge kinds of research.

NSF HERD rankings. Another set of numbers that show up in December for the previous calendar year is the NSF higher education, research and development survey (NSF HERD) numbers. This is where the NSF asks all big schools to respond to a very detailed survey regarding spending of external funding (e.g., Gates, Microsoft, NSF, NIH, DOE, etc.). This is a number that Pitt tracks and reports. The numbers going back 10 years are also published. This is the best measure we have of total Pitt research intensity and the size of the Pitt research enterprise. Pitt fell from 15<sup>th</sup> ranking last year to 18<sup>th</sup> ranking this year. Last year Pitt made a lot of noise in its annual report about having exceeded \$1 billion (\$1.07 billion to be exact) in spending of external research funding. The HERD number is always bigger than the simple sum of sponsored research awards for a given year because it also includes expenditures for building the infrastructure to do the research for which external funding was awarded: facilities, lab renovations, cost-sharing, seed funding, some aspects of start-up packages, and some research training. Since the overall funding number for Pitt exceeded \$1 billion this year, by the NSF HERD formula, Pitt research expenditures will probably amount to \$1.2 billion in the December 2023 report.

### *Questions/Answers*

Z Xia: In the calculation of NIH funding for institutions per fiscal year, do NIH numbers include the funding for the entire project period, or just the portion for that fiscal year? Rutenbar: It's the whole shot. M Holland: Per Tom Berkhoudt in SVC/HS, whose team does the NIH ranking analysis for Anantha: "The amount in the NIH rankings is the award amount given in the NOA for that fiscal year. Sometimes the NIH does fund a project for multiple years in one FFY, and if they do it this way all the funds are counted in the fiscal year the project is awarded. However, most to the time the NIH only funds one year at a time and that is the amount we count."

D Reed: Do NIH numbers include all monies - grants and contracts? M Holland: Real time updating of NIH funding dollars occurs on NIH Reporter and the database locks at the end of the fiscal year (September). The actual final numbers, which include contract funding, lock in in February.

K Wood: How do increases in the cost of doing research contribute to the increases in NIH funding \$\$ year over year for any given institution? Rutenbar: Everybody's numbers are floating up because things are becoming more expensive (inflation) and the NIH builds into the funding increases in salary related expenses, but Pitt's jump from number 6 to number 3 is not due to just inflation alone.

K Wood: Is there a case where regional location within the US makes it cheaper to do research (sourcing supplies, etc.) or is it pretty much the same across the country? Rutenbar: Good

question but I don't have good data on that. M Holland: The Organization of Economic Corporations Development (OECD) does global comparisons of R&D activity and they adjust exactly for that question with a complicated analysis that adjusts for purchasing power parity. That's how you look at that. While there are regional differences in cost of doing research, federal funds presume spending in the U.S. But a primary site of performance overseas generally requires prior permission. Corporations with their own private funding do move R&D across international boundaries all the time. Additionally, F&A rates account, in part, for cost variation across the country. Each university negotiates that rate with their cognizant agency. Pitt's cognizant agency is HHS since it receives more money from NIH than from the DoD (the other cognizant agency for setting F&A rates)

K Wood: Like the movie industry, are there tax breaks given for conducting research in different parts of the US? M Holland: Research and experimentation (R&E) tax credits are sometimes offered by the Federal government and are periodically renewed and periodically lapse. They are a frequent issue of contention. It is argued in Washington/Congress, that unless you make the incentive permanent, which is problematic with respect to scoring issues for budget and deficit purposes, the R&E tax credit really doesn't actually influence behavior. It's a complicated mess that rarely accomplishes what people think it accomplishes.

P Morel: Does this number include training grants? Rutenbar: Yes.

P Morel: Do the NIH numbers include indirect costs in the total costs? Rutenbar: Yes; salary, equipment, travel, fringe benefits, tuition waivers: all that stuff is included in the bottom-line number.

Z Xia: Does the total amount considered for ranking include amounts the primary institution has to subcontract out. Rutenbar: No. It's good to be the prime for accounting purposes.

### **Updates on R3 and data sharing services**

Jonathan Silverstein, MD, MS, FACS, FACMI, Chief Research Informatics Officer, Department of Biomedical Informatics

Uduak Ndoh, MBA, Vice Chancellor and Deputy Chief Information Officer, Health Sciences

There is a strong desire for both Pitt and UPMC to make clinical data available for research. R3 serves that purpose, providing both a service and authorization process for use of multiple data for research. This was structured with adjustments to UPMC policy and legal work between the University and UPMC to produce a business associate's agreement. There are a tremendous number of sources of data images, text, and discrete data for patients that is available for screening. There are multiple layers of technology, data provisioning that go on in generating the data for what can be very large programs. The biggest users are those generating the data.

The request query process. Each request starts with a very brief submission (online form) by the investigator, followed by a conversation with people at R3. The service is billed, all the regulatory issues are addressed, and then the data is delivered in compliance with an IRB approval. Requests that are limited to counts to prepare research do not require IRB approval. There have already been 560 requests in the last 4+ years for which data has been produced.

Cost of services. The scale is driven within the Department of Bioinformatics, which is highly collaborative. Besides the actual core services that are free, there are much bigger projects that have multiple year chargebacks that are scheduled-in yearly charges. There are many collaborative projects that charge for percent of time. The R3 research team includes about 44 people, with a \$5 million range. Chargebacks are about \$300,000 per year.

R3 only has UPMC data. It's a local version of research provided as a service to other investigators. The entire enterprise is based on a set of rules that prioritize the community value of data generated from health care for research purposes and the public good under settings of reduced identification. R3 supports every kind of thing. Most of the work is done with de-identified data – called limited data. R3 supports investigators who are looking for specific identified data under consent of individual patients. Each one has an IRB review of what will precisely happen before any of the research is connected.

Patient registries. R3 often participates in the creation of patient registries for medical researchers, listing hundreds of thousands of lab tests. R3 also supports the registries of individual investigators but they are encouraged not to set up separate registries that are already available through R3. It is a complicated process, reconciling identities into a research identifier that is then deidentified. Alignment of different medical record numbers coming from many different systems/sources is also necessary for bringing the data up to research quality.

R3 as a core service. R3 is supported by the Department of Biomedical Informatics. R3 is open to everyone and is supported by grants from within and by a contract model. There are close collaborations between people at R3, Uduak Ndoh, IRB, CTSI in the SOM Dean's office. The sign-off for the legislative part of it because it's a HIPAA business agreement, occurs between Pitt and UPMC with special privileges for R3 that other honest brokers at Pitt do not have, and a policy structure at UPMC; that was a negotiated process among the legal departments of the 2 universities and the chief medical information officer of UPMC and Dr. Rutenbar's office. There is quite a bit underpinning all of this from a legal structure point of view that was developed over time and would be very hard to replicate. The institutions are required to have things like separation of concerns between clinical data and insurance data, and R3 has special permission to link them together explicitly for research purposes, which UPMC cannot.

### *Questions/Answers*

Z Xia: What is funding the core service? J Silverstein: There have been certain proposals to health sciences, activities with UPMC enterprises wanting to fund this, but no success having direct institutional support despite the institutional nature of what we do. Maybe the SRC group could help with that. Not sure where the money would come from. R3 is running on a modest scale under probably a half dozen grants that include some component of what R3 does (CTSI, tools and technology built specifically for some large common fund programs in biology, that are also usable against these things – such as the Hubmap program translator for building codes). A lot of creativity has gone into getting R3 included into many grants that fundamentally collect data for research; the other half is about charge backs. Because of publicity, we get a little more than we can handle. Some institutional support for R3 would give

it some stability, provide data scientists, and make it easier for people to use for grant-writing. There is no dedicated infrastructure for R3, it takes a little from here and a little from there.

Z Xia: How do you go about soliciting feedback from people who use R3? J Silverstein: We've used the net promoter score approach, which is familiar to folks, to allow people to evaluate the services. The comments we were getting were things that we could not support with what we have: having data scientists to work with them on their projects, or to have certain turn-around times be shorter, when hold-ups are really coming from IRB or grant requirements. There is frustration with wanting to proceed with research without a grant and not having the ability to pay the nominal fees. There are issues around IRB processing such as insufficient experience with the IRB things, where the data will be stored exactly, and having data scientists that work on their projects. So, there's an entire affiliated ecosystem and enterprise that we can criticize and point at implicit opportunities for improvement. I think CTSI does some of that. We try to do what we can. There is vast opportunity for improvement. We don't have the formal oversight that goes with institutional funding, so we have not needed to do it on a systematic, regular basis. It was done as needed. We get regular and continuous feedback because we meet face-to-face with people through this process. We have made use of the feedback.

Z Xia: It sounds like a core service to the university and health scientists, which it is not. J Silverstein: There are a set of core facilities for health sciences, and how you charge and put forth the budget and the business plan, as a department we submit that every 2 years to health sciences. They review our business plan, our budget, and they account for money coming in from grants as baked-in indirect costs - modified total direct costs. To collect \$2K coming from 150 different sources, we must have an infrastructure. We rely on the health sciences core facility structure for their business oversight, but it's purely business-oriented financial oversight, not the oversight that comes with funding or scientific oversight of review that could bring value. Scientific review is something that we have not been required to do; we built the service year over year. Money must come from somewhere.

P Morel: Do you do any kind of data managing? Chart recording, electronic records are not always accurate in terms of codes. Is any type of data curation done? The data can be very messy. J Silverstein: We collect both the structured and unstructured data or text. The structured data is how the system operates, so it is work-flow oriented (never wrong because it's the workflow of what was entered). Whenever you start to use those business-oriented codes for something they were not previously used for (i.e., answering a clinical question), we encourage that there be domain experts and clinicians in the protocol that are validating phenotypes against reading the charts before you scale them up to millions so that you know they are reflecting what you want. Those IRB protocols are more complicated because they must involve clinicians looking at identified data, in some cases without the permission of the patients. The approach is regulated by HIPAA; persons practicing within the UPMC system do have some special privileges regarding the research they are permitted to do.

P Morel: A lot of what goes into the electronic health records is for insurance purposes instead of health care based, which creates a lot of data, a needle in a haystack situation when you haven't previously looked for it. J Silverstein: The retrospective data-oriented projects are equal in thinking (similar logistics) to doing a major clinical trial. Some like to skip over all of that thinking that goes on when you write a trial. We try to encourage not skipping that step but that is not our role in data-provisioning.

S Sant: Is there an update on text searches, which were a struggle pre-Covid? J Silverstein: The text search is much improved; under the right circumstances, it's now instantaneous. Elastic Search, a different kind of database, contains over 400 million individual notes in free text, and its fully indexed. The challenge is the inability to deidentify text to the limited data set, a challenge that the industry has been tackling for some time, reaching 99.5%, but 'federal law requires 100%, an impossibility for all computer science kinds of problems. We're kind of stuck in this funny federal circumstance where the users of that system are the R3 honest brokers. We must sit shoulder to shoulder, often with people on screen, and search within the text for patterns of things of interest to select those patient cohorts and then the patients. Pitt has a fully baked system, the entire structure of which is based in research; its' all there but you can't have it unless you're sitting next to us. It's not our favorite answer, but the one we have today.

S Sant: Most of the people I know who want to do critical translational research are not clinical scientists. Do you provide personnel support through the R3 mechanism? If I tell you the exact question? J Silverstein: The short answer is NO. The long answer is we're situated in the Department of Bioinformatics; we have the people at our fingertips. If R3 were to provide that type of thing, it would be some type of the librarianship of sitting with you and using the tools in a way that is helpful to you. It would be some type of data science-ship, providing people that could do some of the analysis, help with how to navigate, how to make retrospective data useable, a little bit more tool-building to make things more efficient so we can hand you a database rather than get you viewing CSV files. There are many opportunities for improvement. That all requires personnel, and they are expensive. We would need institutional support for that, which we would welcome, but don't yet have.

Z Xia: What do you think is the biggest challenge for R3? J Silverstein: The fundamental question is we are now warming up, almost 5 years since the first days, we're supporting about 150 investigator teams with data and about 5 times that many with administrative interaction, oversight of letters that these R3 requests involve. That's quite a good thing. We've managed to support this thing and grow it with the grants we have and the chargebacks. To go to the next level and provide some of the robust services we have been talking about here would really require us to either charge significantly more for the service, which will make some people go away, or gain some institutional support to provide these kinds of personnel, or people must be extremely patient and write these things into their grants. That means years before you get the opportunity to have those personnel. All of that is happening, except for increasing the cost and providing institution support. The challenge is "do we choose to grow and provide a more robust service for all the things that people are interested in, or do we choose not to? For the SRC to think about that and press different parts of the institution to help with that would be

welcome. We built up quite a bit and at some point, we became an institutional activity, but without institutional support or oversight.

P Morel: That might be a problem experienced by other cores run at the department level but without any institutional support. (Example: Immunology's transgenic mouse/mutant core, which is widely used throughout the health sciences). U Ndoh: Institutional support either through the university or involving the health sciences is something I have been looking at every day, something that Anantha is very aware of, especially as we are looking for more collaborations and more connections. Anantha is very passionate about it; there is "want to support" but it comes down to who is going to cut the check and what do we have to change and prioritize to make that happen.

S Sant: Having that clinical data, especially for getting NIH funding, if you are proposing something out of the box, people really want to see those kinds of associations. We have the resources with UPMC and these other platforms; bringing all that together would be very helpful. In terms of rankings, Pitt can shoot higher (referencing Dr. Rutenbar's presentation). To be able to find associations, identify patients and get the samples to generate more preliminary data from that... It is a missed opportunity for us. If we can get institutional level commitment, then we can come up with some resources. Day in and day out, how many people struggle to work with the data to get that important piece of information that will take their grant to the next level? It is my own experience and that of some of my colleagues. To be able to get institutional support would be really helpful.

Z Xia: Bringing all these resources and backgrounds together will be really helpful. There are missed opportunities for generating preparatory data. Many people like me struggle to work with very important pieces of information. R Rutenbar: Complicated stuff. A conversation that frequently happens at the institutional scale is the need for better support for research cores, things that are critical for getting next-gen research launched, make it competitive and great, research that is really big and really complicated. There is no individual faculty or department that can afford this and it needs to be more widely supported. There was historically kind of a central model/mechanism for that stuff. We are having the conversations from different departments about the scholarship – the same problem. We have historically been so decentralized about this, there's a point at which you want to guard - a little organization. Not sure what the right answer is, but we are having the conversations. You don't have to convince me.

Z Xia (addressing R Rutenbar): It is your job to go and convince the others. P Morel: Invest in this and we get the \$10 million. You think this is something that UPMC Enterprises would want. U Ndoh: More UPMC investment in research would benefit them in the long run, come back to them (bench to bedside), but it hasn't worked out for them that way yet.

**New items of discussion: none**

**The meeting was adjourned at 2:45pm.**

**The next Research Committee meeting:** March 17 (unless issues arise in the meantime)

**Minutes submitted by:** K Wood and M Scott, with edits from M Holland and Z Xia.