Research Infrastructure for Collaborative Team Science: Challenges in Technology-Supported Workflows in and Across Laboratories, Institutions, and Geographies

Barbara Mirel, DArts,[*] Airong Luo, PhD,[†] and Marcelline Harris, PhD[‡]

Summary: Collaborative research has many challenges. One under-researched challenge is how to align collaborators' research practices and evolving analytical reasoning with technologies and configurations of technologies that best support them. The goal of such alignment is to enhance collaborative problem solving capabilities in research. Toward this end, we draw on our own research and a synthesis of the literature to characterize the workflow of collaborating scientists in systems-level renal disease research. We describe the various phases of a hypothetical workflow among diverse collaborators within and across laboratories, extending from their primary analysis through secondary analysis. For each phase, we highlight required technology supports, and. At time, complementary organizational supports. This survey of supports matching collaborators' analysis practices and needs in research projects to technological support is preliminary, aimed ultimately at developing a research capability framework that can help scientists and technologists mutually understand workflows and technologies that can help enable and enhance them.
Semin Nephrol 35:291-302
Keyword: Systems biology, bioinformatics, transciplinary research, interdisciplinary research, research capability model, technological innovation

The Nephrotic Syndrome Study Network encourages collaborations for integrative biomedical research across disciplines, subspecialties, institutions, and geographies. Working together, researchers access, manipulate, transform, validate, and share data and knowledge for novel insights into molecular mechanisms of renal disease. With combined expertise they discover and confirm what none alone could have done. However, collaborative research has many challenges. Some challenges have been examined closely by team science researchers, and they are related mainly to interpersonal and organizational dynamics for collaborative readiness.[1] Other challenges, however, are not so well researched, especially the challenge of achieving seamless flows of technology-enabled, discovery-driven analyses within and across collaborating groups. A seamless flow of analysis calls for aligning collaborators' research practices and evolving analytic reasoning to the technologies that best support them.[2] Supportive technologies are wide-ranging and may include databases and data management tools, security protocols, high-throughput instrumentation, diverse software applications, algorithms, data transfer protocols, ontologies, and web resources and services. These technologies together with equipment, services, computational resources, and domain tools constitute a research infrastructure. In one survey, team scientists ranked adequate and appropriate resources and infrastructure as a top 10 need for productive research.[3]

Importantly, a research infrastructure involves more than the availability of enabling technologies and services. It requires combinations and configurations of them that will accomplish the "ultimate goal [of] … allow[ing] scientists to enhance their collaborative problem solving capabilities through the improved and integrated usage of resources and tools".[4] A research infrastructure implements requirements for research capabilities. We define research capabilities as competencies for leveraging human, organizational, and technical resources and services for purposes defined by the goals of a research project. Translational researchers note that when resources and infrastructures are insufficiently matched and configured to their needs and purposes, their research progress tends to be delayed. Moreover, they often have to re-invent the wheel in each project in terms of logistics, data exchanges, harmonization, databases, and interfaces. This fitness-to-purpose hinges on aligning technologies with analysts' reasoning and behaviors, which, in turn, requires a good understanding of researchers' goal-driven workflows and the challenges they encounter in them for their analytic needs.[4–6]

In this article, we seek to advance this understanding. We describe a hypothetical workflow for integrative

[*]*School of Education, University of Michigan, Ann Arbor, MI.*
[†]*MSIS Research IT, Medical School, University of Michigan, Ann Arbor, MI.*
[‡]*School of Nursing, University of Michigan, Ann Arbor, MI.*
Financial support: Supported by National Institutes of Health (1 P30 DK081943-01).
Conflict of interest statement: none.
Address reprint requests to Barbara Mirel, DArts, University of Michigan, 610 South University, 2602 School of Education Building, Ann Arbor, Michigan 48108. E-mail: bmirel@umich.edu
0270-9295/ - see front matter

http://dx.doi.org/10.1016/j.semnephrol.2015.04.009

renal disease research. For each phase of the workflow we describe associated challenges. In practice, some challenges may recur across phases but for purposes of analysis we tie them to the phase in which they are most prominent. For the challenges, we propose technologic supports and, at times, complementary organizational supports that may enhance researchers' capabilities to address them. We categorize these supports by the type of research infrastructure requirement they connote. We strive to frame supports and categories in terms that will resonate with technology and organizational stakeholders. Toward this end, we adapt the language used in well-established capability maturity models and frameworks.[5,7,8] Capability maturity models and frameworks address processes and resources that organizations and information technology units need to provide to meet business requirements.

Our adaptive uses of capability maturity models' terms and categories for collaborative research infrastructures are distinct. To our knowledge, little research centers, as we do, on the perspective of collaborating researchers' flow of integrative biomedical analyses to identify unified sets of support for this research. From this perspective, we uncover combinations of technologies and organizational processes that need to be well-integrated to mitigate challenges that renal disease researchers may encounter in their systems-level analytic workflows. As a caveat, we do not provide "how-to advice" (eg, specific tool recommendations or configuration designs for solutions).

Our framing is preliminary and currently ongoing. By applying it here, we hope to help collaborating renal disease researchers recognize various connections between their workflows, technologic challenges, necessary supports, and categories of support in the research infrastructure. With this awareness, collaborating researchers may be better able to pre-emptively plan for and address these challenges. As neuroscience researchers have found in an outcome that likely is relevant to renal disease, "the growing importance of a complex, interoperable [information technology] IT-based research infrastructure is underestimated in many research designs and could be optimized."[9] With this awareness, researchers also may be better able to articulate and explain their research requirements to information technology (IT) units and together negotiate services and resources that enable collaborative research.

OVERVIEW: RESEARCH WORKFLOW AND CHALLENGES

To construct the hypothetical workflow we synthesized studies from the research literature related to team science, computer-supported collaborative works, and analytic and visual analytic workflows in -omics inquiries.[1–3,10–21] We combined this synthesis with our own prior research on team science workflows and collaborations.[22–27] The workflow is generalized based on common patterns found in the research literature. Inescapably, workflows for systems-level

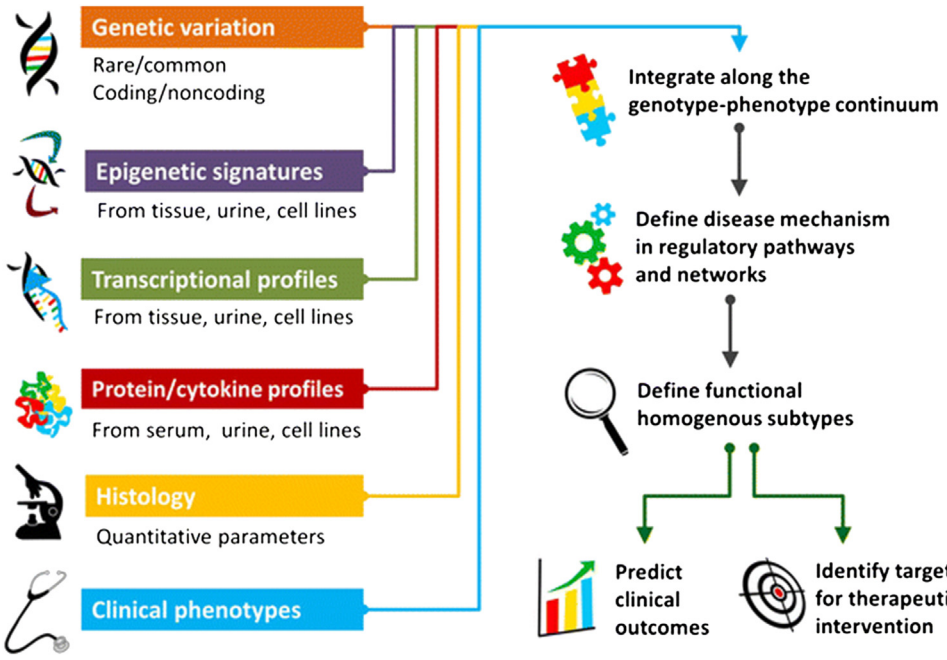

Figure 1. Multiple biological scales feeding into integrative renal disease research. Reprinted with permission from Keller et al.[29]