National Institutes of Health National Heart, Lung, and Blood Institute (NHLBI)

Final Program Evaluation Report

Contract No. GS10F0088P/HHSN268201400126U

April 2016

Submitted to:
Dr. Pothur Srinivas, PhD
Division of Heart and Vascular Disease
NHLBI/National Institutes of Health

Mona Puggal, MPH
Division of Cardiovascular Sciences
NHLBI/National Institutes of Health

Reported by:
Dr. Scott Rosas, PhD
Dr. Sherri Brokopp-Binder, PhD
Jessica Gosa, MSW, LMSW
Jennifer Royer, MPA
Colleen Budzinski, MS
Steven Vedder, BA
Concept Systems, Incorporated

Table of Contents

I. Executive Summary

The full report describes the evaluation of the National Heart, Lung, and Blood Institute (NHLBI) Proteomics Centers Program established in 2010 through the National Institutes of Health (NIH). This evaluation was funded by the NIH Evaluation Set-Aside Program, (GS10F0088P/HHSN268201400126U), administered by the Office of Program Evaluation and Performance, DPCPSI, Office of the Director. The full report is an assessment of the NHLBI Proteomic Centers Program progress in meeting its short- and long-term objectives to-date. Concept Systems, Inc. (CSI) was hired as an external evaluator to conduct an evaluation of the Program in the final iteration of the grant cycle from 2010-2015. This Executive Summary provides a general overview of the Program, evaluation design, and selected findings.

Background

The goal of the external evaluation is to help determine whether milestones are being met and to help identify potential pitfalls and strategies moving forward.

The Proteomics Centers Program consisted of the seven contracted NHLBI Research Centers and an Administrative Coordinating Center (located at the University of California Los Angles):

1. Boston University Cardiovascular Proteomics Center
2. Harvard-Broad Proteomics Center
3. John Hopkins University Proteomic Innovation Center in Heart Failure
4. Stanford University Proteomics Center
5. University of California Los Angeles Proteomics Center- Global Proteomic Initiative of Cardiovascular Medicine
6. University of Texas Health Center at San Antonio Cardiovascular Proteomics Center
7. University of Texas Medical Branch NHLBI Proteomics Center at Galveston

The primary objective and purpose of using a center-based science model to advance the knowledge and discovery of proteomic research was to integrate multidisciplinary methods of research and collaboration specific to the areas of heart, lung, and blood. The model was characterized by the integration of three essential components: 1) proteomic technology development, 2) mechanistic and functional proteomic studies, and 3) proteomic clinical applications. The findings from CSI's evaluation that are found in this report are based on those key activities, as assessed through a data collection process described in the Evaluation Data Source Matrix (**See Appendix B. Evaluation Data Matrix**).

Evaluation Design

In order to gain a robust understanding of the Program and its impacts, CSI used a mixed-method evaluation design with both qualitative and quantitative approaches to data collection and analysis. In addition to the evaluation findings, the full report includes a discussion of overall observations and conclusions related to the NHLBI Proteomic Centers Program. Supporting information, figures, tables, and appendices are provided throughout the full report.

The evaluation of the NHLBI Proteomic Centers Program ultimately sought to answer the global question of, "*what difference did the Program make*"? Within this frame, this evaluation articulated seven questions to assess and examine the extent to which the Program accomplished the intended goals. These questions were developed based on the objective outlined in the NHLBI request for qualifications (RFQ) and input from an Expert Advisory Panel (**See Appendix A. Expert Advisory Panel**) assembled by the evaluation team. The seven questions driving this evaluation included:

1. To what extent has the Program developed and shared new tools and technologies?

2. What outputs of the Program have matured enough to be of clinical utility? To what extent has the Program resulted in inventions, patent applications, spin-off companies, and licensing agreements?
3. To what extent has the research contributed to the creation and integration of a knowledgebase linking changes in proteomes with molecular phenotypes of disease?
4. In what areas has the Program filled existing knowledge gaps in heart, lung, and blood disease biology?
5. To what extent have the Program investigators been collaborative within and outside their fields?
6. To what extent was subsequent science informed by the research conducted by the Program?
7. To what extent has the Program served as a springboard for subsequent academic appointments and professional recognition?

Selected Findings

The Program was formed around the creation of multiple integrated Centers, each was tasked with forming interdisciplinary teams of investigators and trainees to address a suite of research questions. It was expected that these teams, in turn, would support the development of a shared database and foster an increase in shared understanding of shared tools and knowledge.

The Extent the Program Developed and Shared New Tools and Technologies

The Centers engaged in multiple efforts to expand the use and appreciation of proteomics by the broader scientific community globally. Respondents described several new tools and technologies developed as part of the Centers' scientific agenda. Respondents also highlighted several recognizable advancements, and how the Program allowed new tools to be vetted and improved. As a significant resource for the field of proteomics, the CoPaKB database was created by the Program as a unique resource to facilitate discovery of novel biological insights from proteomics datasets. The web-based CoPaKB portal went live in the Program's second year and access increased steadily each year. The presence of the NHLBI Center program scientific output in the hierarchy of peer-review journals also speaks to the visibility and credibility, which is described in detail in the full report.

The Extent the Program Created Outputs of Clinical Utility

Investigators within the Program reported having made significant progress in the development of tools and generation of knowledge with clinical relevance; several examples of clinical significance of their accomplishments were described. At the same time, it was clear from the data that the advancements towards clinical relevance and clinical applications was a work in progress. The assessments of respondents who were external to the Program were equivocal; some were in agreement that the Program had made significant progress in this areas, others were much more guarded and skeptical in their assessments.

The Extent Research Contributed to Knowledgebase Linkages

There was general agreement among respondents that over the course of the Program, significant growth has occurred in the field. The field has progressed in technical development, engineering research, and clinical advancement. Investigators at most Centers reported having confirmed linkages between specific proteomes and the molecular phenotypes of the disease or diseases in which they were focused. However, there are differing opinions, particularly by individuals internal and external to the Program, about the extent to which the Program has contributed to these advances.

Areas the Program Filled Knowledge Gaps in Heart, Lung, and Blood Disease Biology

Respondents described the research goals of their Centers and labs, including goals related to technological advancements, basic science research questions, and clinical goals. While most respondents were able to articulate these goals, some trainees and investigators struggled to state the overall goals of their Center. Network analysis findings of the topical network of scientific output produced by the Centers suggest confirmation at a high level that the areas of the scientific inquiry set forth by the Centers individually, and the Program collectively at inception, were manifest in the scientific output produced during the period.

The Extent the Program Investigators Collaborated Within and Outside the Field

In terms of the patterns of collaboration based on co-authorships, the network analysis detailed in the full report revealed the growth of an extensive collaborative network over the period of the Program. As the network accumulated, more researchers engaged in co-authored studies and changes to both the size and cohesiveness of the network was noted. From the interviews, respondents described many ways in which they collaborated with other researchers across Centers and outside of the Program, globally.

The Extent Science Informed by the Research was Conducted by the Program

Overall, investigators within the Program reported having made significant progress in the development of tools and generation of knowledge with clinical relevance, and provided examples of the clinical significance of their accomplishments. However, there were some differing opinions, particularly by individuals internal and external to the Program, about the extent to which the Program has contributed to these advances. The bibliometric analyses indicates that the NHLBI Proteomic Centers research is generating substantial interest.

The Extent the Program has been a Springboard for Subsequent Academic and Professional Recognition

Overall, the Program fostered an environment within which many stakeholders felt supported in the development of their careers. While interview respondents commented on how opportunities to collaborate outside of their Centers had contributed to their training, there were also some hindrances shared as well. With examination of the biosketches and curriculum vitaes of the Centers' key personnel, several instances of position advancement occurred during the period of the Program. Additionally, findings suggest that individuals affiliated with the Program were recognized both nationally and internationally. Findings related to scientific recognition on the basis of citation patterns suggest cumulative scientific output of the Centers to be substantively recognized by other scientists.

Overview of Findings and Conclusions

Ultimately, the findings of this evaluation help to support and further explain the NHLBI Proteomic Centers Program's innovative center-based outreach, research, and collaboration. Overall, stakeholders recognized the Program's expanded capacity to share tools, knowledge, and further engage and develop proteomic research. Furthermore, this group commented on and acknowledged that the Program contributed to a network of collaborations and partnerships, facilitated the training of promising early-career professionals, leveraged funding, established physical and virtual laboratories, and generated data clearinghouses. In this evaluation, we found evidence of highly recognized and scientifically informative corpus of scientific output. The collective body of scientific work, as represented by the substantive volume of peer-reviewed papers, was performing at a high level compared to what was expected based on where and when the work was published.

II. Introduction

Background

The National Heart, Lung, and Blood Institute (NHLBI) Proteomics Centers Program was established in 2010 with the goal of applying proteomic approaches to gain a better mechanistic understanding of the physiologic pathways underlying defined clinical conditions related to heart, lung, and blood diseases. The Program was administered within the Division of Cardiovascular Sciences at the NHLBI from August 2010 through August 2015. The mission of NHLBI is to:

> *"…provide global leadership for research training, and education programs to promote the prevention and treatment of heart, lung, and blood diseases and enhance the health of all individuals so that they can live longer and more fulfilling lives".*

In establishing the Proteomics Centers Program, NHLBI intended to support the development and discovery of new tools, proteomic applications and a knowledgebase to facilitate proteomic clinical approaches and efficacy.

The Proteomics Centers Program consisted of seven contracted research Centers[1] and an Administrative Coordinating Center[2]. The individual Centers worked to integrate multidisciplinary expertise from fields that included proteomics, physiology, clinical studies, molecular biology, genomics, chemistry, physics, engineering, computational biology, bioinformatics, and biostatistics to advance proteomic applications in heart, lung, blood, and sleep diseases and disorders. Each NHLBI Proteomic Center identified and addressed a specific clinical need/problem/disease/process integrating three components to form an interactive team:

1. proteomic technology development;
2. mechanistic and functional understanding of the proteome, its interactions and dynamics; and
3. clinical application of proteomic approaches and discoveries.

The components were expected to be interdependent and drive each other's science and progress.

NHLBI Proteomics Centers Program Purpose and Goals

The NHLBI Proteomics Centers Program was administered through a five-year contract with individual Centers for the purpose of establishing interactive, multi-disciplinary teams of experts in scientific diverse areas. The Centers, separately and collectively, were expected to enhance and develop innovative proteomic technologies, and apply them to biological questions relevant to NHLBI. Specific areas addressed included heart, lung, blood, sleep, chemistry, physics, engineering, proteomics, informatics and statistics. The design of the Program was intended to foster a greater understanding of physiological pathways, molecular interactions, and regulatory signals related to heart, lung, blood and sleep diseases and disorders, as well as to begin to apply the technologies and knowledge to relevant clinical questions. Innovation and creation of new technologies and tools was emphasized with the intention that advancements would be made available to the community at-large for further advancement of heart, lung, and blood research. Training and research opportunities for both senior and junior investigators were also supported by the Program.

The NHLBI Proteomics Centers Program was consistent with the assumptions and expectations of collaborative center-based science. Briefly, such collaborative center-based models are growing in

[1] Boston University Cardiovascular Proteomics Center; Johns Hopkins University Proteomic Innovation Center in Heart Failure; Harvard-Broad Proteomics Center; UT Health Center at San Antonio Cardiovascular Proteomics Center; Stanford University Proteomics Center; University of Texas Medical Branch NHLBI Proteomics Center at Galveston; UCLA Proteomics Center- Global Proteomic Initiative of Cardiovascular Medicine

[2] NHLBI Proteomics Coordinating and Administrative Center - UCLA

frequency, size, diversifying in scope, and are foundational to an increasing effort to expedite scientific inquiry and discovery. Many of the distributed and collaborative scientific research centers are based on the belief that team-based research integrating the strengths of multiple disciplines may accelerate progress toward resolving complex societal and scientific problems.[3,4] Dispersed across different departments, institutions, and geographic locations these collaborative centers are intended to exploit opportunities in science, engineering, and technology in which the complexity of the research problem or the resources needed to solve the problem require the advantages of scope, scale, duration, equipment, facilities, and students.[5,6,7]

Through the integrated approach, each NHLBI Proteomic Center was expected to enhance proteomic tools and technologies with the goal of using them for mechanistic and functional understanding of heart, lung, blood, and sleep diseases and disorders. Enhancement of proteomic tools and technologies could include hardware or software to improve the measurement, characterization, or comparison of proteins, peptides, PTMs, isoforms, genomic/epigenomic tags, small molecules, or other related cellular or tissue components. The mechanistic and functional understanding could be focused on experimental and/or computational approaches to molecular function/dysfunction, dynamics, interactions, pathways, process, transport, or other related actions, in cells, tissues, organs, and whole organisms. The Centers were to work toward application in clinical studies using the proteomic tools, technologies, mechanisms, and functional understanding to guide human subject studies. It was expected that all Centers would have human research studies during the course of the project period. The Centers were to leverage existing resources to support these human subject studies, such as CTSA, clinical networks, ongoing clinical trials, ongoing cohort studies, and industry supported studies. It was also expected that the funded Centers would interact to share information on technical objectives, progress and impediments, as well as exchange ideas, and, where appropriate, establish collaborations. The Centers were directed to make available to the research community all research products and results in a timely fashion. Each Center developed a project plan that addressed milestones, deliverables, and sharing.

Program Funding Mechanism

The NHLBI Proteomics Centers Program was unique in that individual Centers were selected for awards through a Broad Agency Announcement (BAA). Pertaining to basic and applied research, BAAs are diverse in their subject matter and focus on advancing science rather that acquiring specific products. The government agency issuing BAAs provide research and technical objectives the agency is interested in, rather than an established scope of work. The offeror, not the agency, develops the scope of work including the requirements and performance specifications. A contracting relationship is then established to negotiate, refine and incorporate peer-review and program comments. In this regard, the use of and expectations for contracts as a funding vehicle are different from that of grants or cooperative agreements. This mechanism was employed to encourage the submission of creative innovative methods with varying technical/scientific approaches. While the Centers were expected to incorporate core elements as outlined by NHLBI and the contract, there was a degree of flexibility afforded for Centers to institute diverse scientific agendas.

The Program was funded from August 2010 through August 2015, and had a total cumulative budget of approximately $83.5 million dollars. Individual Centers were awarded varying funding amounts depending on the scope of their work. There was a previous iteration of the Program from September

[3] Borner K, Contractor N, Falk-Krzesinski HJ, et al. (2010). A multi-level perspective for the science of team science. *Sci Transl Med*. 2(49).
[4] Crow MM (2010). Organizing teaching and research to address the grand challenges of sustainable development. *BioScience*. 60(7):488–9
[5] National Cancer Institute: *Science of team science*.
[6] Olson GM, Zimmerman A, Bos N (2008). *Scientific collaboration on the Internet*. Cambridge, MA: MIT Press.
[7] Trochim WM, Marcus S, Masse LC et al. (2008). The evaluation of large research initiatives: a participatory integrated mixed-methods approach. *Am J Eval*. 29:8–28.

2002 to September 2009, which included a consortium of 10 multidisciplinary Proteomic Centers. This previous Program iteration focused intensively on the development of proteomic technologies.[8]

NHLBI Proteomics Centers

Seven institutions were awarded funding for the 2010-2015 iteration of the NHLBI Proteomics Centers Program. Three awardee institutions were new and not part of the previous program iteration. The Center sites and their clinical focus areas are shown in **Table 1 below**.

Table 1. NHLBI Proteomics Centers Program, and Clinical Focus

NHLBI Proteomic Center	Clinical Focus
Boston University Cardiovascular Proteomics Center[a] Boston, MA Director: Cathy Costello, Ph.D.	Cardiovascular Disease (CVD)
Harvard-Broad Proteomics Center/Massachusetts General Hospital Boston, MA Director: Robert Gerszten, M.D.	Myocardial Ischemia
Johns Hopkins University Proteomic Innovation Center in Heart Failure[a] Baltimore, MD Director: Jennifer Van Eyk, Ph.D.[b]	Heart Failure
Stanford University Proteomics Center[a] Stanford, California Director: Gary Nolan, Ph.D.	Pulmonary arterial hypertension (PAH)
UCLA Proteomics Center- Global Proteomic Initiative of Cardiovascular Medicine Los Angeles, California Director: Peipei Ping, Ph.D.	Cardiovascular Disease
UT Health Science Center at San Antonio Cardiovascular Proteomics Center San Antonio Director: Merry Lindsey, Ph.D.	Cardiovascular Research
University of Texas Medical Branch NHLBI Proteomics Center at Galveston[a] Galveston, TX Director: Alexander Kurosky, Ph.D.	Airway inflammation related to Asthma, Allergies and Respiratory Viruses

[a] Centers that were part of both Program Iterations

[b] Dr. VanEyk changed institutions in 2014 to Cedars Sinai Hospital in Los Angeles

[8] NHLBI Website: http://www.nhlbi.nih.gov/research/resources/proteomics

III. Evaluation Purpose and Approach

Given the rapid developments in proteomic technologies and approaches in the last five years, NHLBI recognized the importance and value of assessing the extent to which these have matured, leading to discovery of new targets for intervention and clinically actionable tool sets. The results of this evaluation were intended to help determine the extent to which the desired outcomes of the Program were achieved as well as to inform the future of proteomics research funding and commitments by the NHLBI.
The NHLBI, Office of Acquisitions (OA), Consolidated Operation Acquisition Center (COAC) Services Branch, provided support for the evaluation of the NHLBI Proteomics Centers Program in accordance with Part III- Statement of Work (SOW). In accordance with the procedures of FAR subpart 8.405-2, the NHLBI awarded Concept Systems, Incorporated (CSI) a single task order to provide an 18-month program evaluation. This evaluation was funded by the NIH Evaluation Set-Aside Program, (GS10F0088P/HHSN268201400126U), administered by the Office of Program Evaluation and Performance, DPCPSI, Office of the Director.

Given the uniqueness of the Centers and the absence of a reasonable comparator, we employed an evaluative framework that sought to answer the global question of, "*what difference did the Program make*"? In order to describe and assess the cumulative success of the Proteomics Centers Program our approach focused on developing:

- A logical explanation for why the investment can be expected to have led to the observed outcomes.
- A plausible time sequence of the investment that occurred and the observed change relative to an appropriate starting point.
- Compelling evidence that the investment/actions are the partial or full cause of the change when competing explanations are taken into account.

The evaluation team at CSI incorporated a contribution analysis[9] approach that aided in constructing a credible explanation of what occurred in the Program and lead to the intended outcomes. A multi-step, iterative process, contribution analysis is often used in complex, multi-level scenarios to examine context, mechanisms, and outcomes to see what worked under what circumstances, and the role the program played in the larger system. Use of the contribution analysis was consistent with the broad evaluation purpose for the Proteomic Centers Program, embraced "plausible association" perspective, and relied upon multiple sources of evidence. In adhering to a contribution analysis perspective the evaluation offered a framework for:

- providing a well-articulated presentation of the context of the Program and its general goals, objectives and activities to achieve those ends;
- presenting a plausible theory of the strategy leading to meeting the overall goals of the Program;
- describing the activities and outputs produced by the Program;
- highlighting the results of the contribution analysis indicating there is an association between what the programmatic strategy has done and the outcomes observed; and
- illuminating the main alternative explanations for the outcomes occurring have been ruled out, or clearly had limited influence.

[9] Mayne J (2001). Addressing attribution through contribution analysis: using performance measures sensibly. *Can J Prog Eval*. 16(1): 1.

Evaluation Questions

Evaluation questions relevant to the NHLBI Proteomics Centers Program were outlined by NHLBI as part of the RFQ. Seven broad evaluation questions were framed to assess and examine the extent to which the Program accomplished intended objectives:

- **Question 1**: To what extent has the Program developed and shared new tools and technologies?
- **Question 2**: What outputs of the Program have matured enough to be of clinical utility? To what extent has the Program resulted in inventions, patent applications, spin-off companies, and licensing agreements?
- **Question 3**: To what extent has the research contributed to the creation and integration of a knowledgebase linking changes in proteomes with molecular phenotypes of disease?
- **Question 4**: In what areas has the Program filled existing knowledge gaps in heart, lung, and blood disease biology?
- **Question 5**: To what extent have the Program investigators been collaborative within and outside their fields?
- **Question 6**: To what extent was subsequent science informed by the research conducted by the Program?
- **Question 7**: To what extent has the Program served as a springboard for subsequent academic appointments and professional recognition?

Expert Advisory Panel

Give the expectation set forth by the NHLBI, CSI formed an evaluation Expert Advisory Panel as part of the overall evaluation approach. Six panelists were selected for their high level of subject matter expertise and are in a recognized position in the topic areas of the evaluation. For the NHLBI Proteomics Centers Program evaluation, panel membership included those that have expertise in proteomics, as well as panelists that have expertise in the field of evaluation.

The Expert Advisory Panel for the NHLBI Proteomics Centers evaluation served as a technical resource to the evaluation team, providing critical guidance to the work so that it is well-informed, efficient, and sensitive to the Program context. From an expert advisory position, the Expert Advisory Panel helped to ensure the evaluation was conducted in a standards-based manner to (1) meet the needs of those who intend to use the evaluation findings; (2) be practical, doable, and realistic; (3) yield findings that are correct; and (4) be conducted with awareness of the rights of stakeholders.

The Expert Advisory Panel provided important feedback on design, tools, analysis, results, and report development. The panel met via conference call four times over the course of the 18-month evaluation, and several members were consulted with directly through email, based on their specific expertise. The names and a short biography of the panel members can be found in **Appendix A: Expert Advisory Panel**.

Management Approach

The CSI evaluation team worked with NHLBI Leadership and NIH associated staff to identify and collect appropriate data elements as well as to obtain and clarify program related information relevant to the evaluation. Input from the NHLBI Project CORs was acquired during all phases of the evaluation through biweekly project planning calls and onsite quarterly meetings.

The evaluation team also employed a planned approach to project management that enabled successful completion of NHLBI Proteomics Centers Program evaluation. These project management methods allowed opportunity for flexibility in meeting project goals, which is necessary in the evaluation of

scientific research systems, where objectives change as new information is ascertained. The project management process included establishing, managing and conducting regular communication with working groups and the advisory panel to refine project goals and scope and contribute to other key decisions.

IV. Methodology

In order to meet the purpose of the outcome evaluation, CSI designed and implemented an *integrated, mixed-methods evaluation approach*. This approach integrated qualitative and quantitative methods to data collection and analyses. Such an approach enabled the evaluation team to (1) enhance the validity or credibility of evaluation findings through results from the different methods that converge and agree, (2) extend the comprehensiveness of evaluation findings through results from different methods that broaden or deepen the understandings reached, (3) generate new insights in evaluation findings through results from different methods that diverge, and (4) incorporate a greater diversity of values about what constitutes programmatic success in the context of internal and external challenges.

The evaluation approach strongly emphasized quality and strived to meet evaluative standards set forth by the evaluation field of accuracy, propriety, feasibility, and utility. For each of the broad evaluation questions, multiple qualitative and quantitative data points were used to yield answers. Integration involved subjective and objective sources of information at several levels, including data collection, analysis, and reporting.

Qualitative and quantitative data were collected on multiple aspects of the NHLBI Proteomics Centers Program including the following: semi-structured interviews, NIH biosketches/curriculum vitaes, Centers' progress reports, scientific research publications, scientific aims of the Centers, conference summaries, web-analytics and altmetrics data. An evaluation data source matrix was developed that identifies specific data elements collected to examine each of the evaluation questions (**See Appendix B: Evaluation Data Source Matrix**).

In each case, the evaluation team developed and utilized formal data collection protocols to ensure consistency, accuracy, and usability. To tie various data elements across the sources together, a relational multi-table database was developed for the qualitative and quantitative data separately. This database provided a structure for organization and management of the data collection to be easily queried to facilitate analysis. Due to the broad range of activities, outputs, and outcomes of complex research initiatives like the NHLBI Proteomics Centers Program, it was essential to use multiple perspectives, methods, and sources as part of the evaluation. Brief summaries are provided below of all quantitative and qualitative data sources, and highlights regarding collection methods. Specific details regarding protocol development, data collection procedures, quality assurance processes, and general analyses are shown in **Appendix C: Key Informant Interviews and Qualitative Analysis** and **Appendix D: Data Abstraction/Scientific Literature Review, Bibliometrics, Altmetrics, and Publication Analysis**.

Evaluation Logic Model

In order to help guide and shape the identification of indicators and data sources in the evaluation, a logic model was developed and refined throughout the evaluation with the Expert Advisory Panelists as well as with the NHLBI Program Officers and CORs. In general, a logic model captures stakeholders' assumptions about how the different resources and activities lead to the desired outcomes and ultimate impact. It describes the presumed program or initiative theory and conveys the sequence of expected processes and outcomes. The logic model maps out and represents the linear sequence that shows how the logic of the program leads from inputs, activities, and outputs to the short-term, intermediate and long-term outcomes. A logic model helps to articulate specific, detailed, measurable and objective program evaluation questions.

The NHLBI Centers Program logic model reveals four major pathways: *Strategize and Coordinate*, *Engage Research and Development Community*, *Interactive Team Research and Training*, and *Disseminate New Tools or Prototypes*. The four major areas on the logic model have been connected with the indicators, data sources, and preliminary collection approaches relevant to the described short-term

outcomes of the Program. The evaluation team developed and tested multiple protocols with procedures that are consistent with each data source as shown **Figure 1 below**.

NHLBI Proteomic Centers Program Logic Model

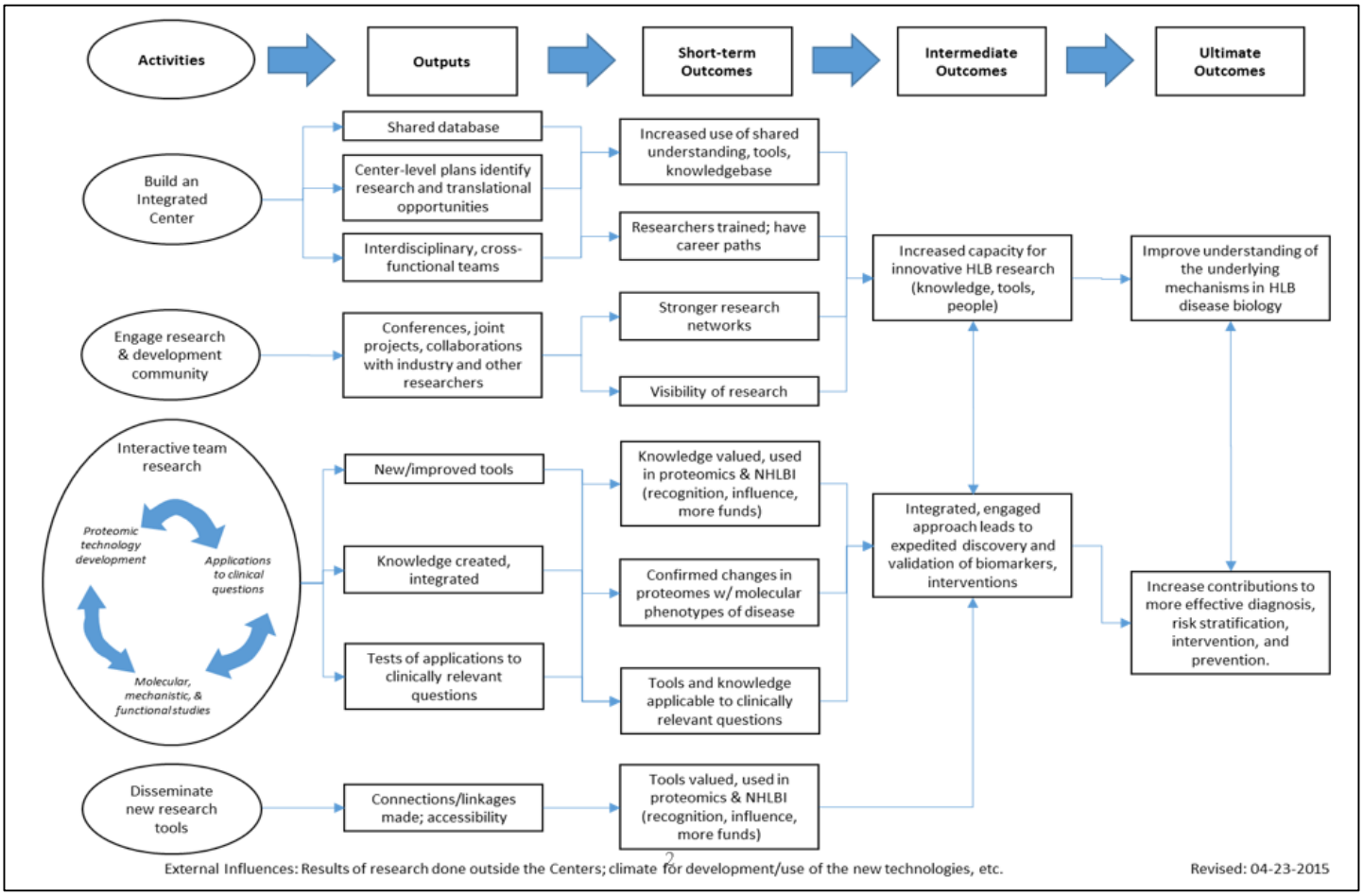


Figure 1. NHLBI Proteomic Centers Program Logic Model

Qualitative Interviews

Subjective input to the success of the Program was captured through semi-structured, qualitative interviews with respondents. A full description of the data collection and analytical methods can be found in **Appendix C: Key Informant Interview and Qualitative Analysis**. Briefly, these key informant interviews were conducted with four targeted groups:

1. principal investigators and key personnel from the seven Centers;
2. proteomics investigators not funded by the NHLBI Program (Outside Experts);
3. trainees (current and former) and junior investigators; and
4. representatives from NIH, NHLBI, and NCI (referred to collectively as NIH) who have knowledge of or who have participated with the Program in some capacity.

Interviews were conducted to facilitate discussion and draw insights from multiple perspectives about the Program, and to yield appropriate information to address the evaluation questions. We assumed individuals from these respective groups had unique perspectives as the successes and limitations of the Program. Therefore, it was critical the semi-structured interview guides were developed in a way that enables individuals to provide information shaped by their experiences, in the context of the overarching

questions. Using the evaluation questions as the guiding framework for interview question development, the interview guides were used to draw insights from multiple perspectives and triangulate sources of information in an effort to confirm observations and conclusions.

A pilot phase was conducted between January 2014 and February 2015, in order to (1) to ensure that the semi-structured interview guides appropriately framed the perspective of the individuals from the four groups; and (2) to obtain data information collection clearance from the Office of Management and Budget (OMB).

Following refinement to the protocols and adjustment to the procedures based on feedback in the pilot phase, the "full-set" of interviews took place between June 2015 and October 2015 with 65 individuals across four groups. This number was broken down by group as follows: 27 Principal Investigators (PIs) (including Center Directors, Co-PIs, and Investigators), 20 Trainees (including both current and former), 9 Outside Experts, and 9 NIH (**Table 2 below**). In order to reach this goal, interviews were conducted both at the NHLBI Proteomics Center PI Meeting in Bethesda, Maryland in June 2015, and remotely by phone or Skype.

Table 2. Interview Respondents by Group

Respondent Group	Target Number	Number Completed
PIs		
Directors	7	7
Co-PIs	14	12
Investigators	6	6
Trainees		
Current	10	9
Former	10	9
Outside Experts	9	9
NIH	9	8
TOTAL	**65**	**60**

Interviewers captured the content of the interviews through digital recording, from which summaries and transcripts were produced. Through an analytic memo process, the interviewers used the interview guide to record insights and summarize reflections after each interview. To ensure the guides were yielding the anticipated information and feedback, this debriefing process was essential throughout the interview data collection. In addition, interviewers followed up with the interview respondents, when necessary, to ensure validity of the data. Upon verifying the transcription with the audio recording, a qualitative data reduction and coding process was initiated. Following an iterative analytic process using multiple analysts on the team to enhance the credibility of the findings, a codebook was created to structure and guide the analysis. Through this process, the qualitative interview data was analyzed, and the thematic results were interpreted in the context of the broad evaluation questions. The use of respondent language was emphasized and verbatim accounts identified to amplify key themes that emerged from the analysis. A complete methodological description of the qualitative interview process, tools, and codebook summary can be found in **Appendix C: Key Informant Interview and Qualitative Analysis.**

Data Abstraction/Scientific Literature Review, Bibliometrics, Altmetrics, and Publication Analysis

Objective input to the success of the Program was captured through the collection, coding, and analysis of NIH biosketches/curriculum vitaes (CVs), Centers' progress reports, scientific research publications, scientific aims of the Centers, conference summaries, web-analytics and altmetrics data. A full description of the data collection and analytical methods can be found in **Appendix D: Data Abstraction/Scientific Literature Review, Bibliometrics, Altmetrics, and Publication Analysis.**

Data Abstraction

For the data requiring abstraction and coding (i.e., NIH biosketches/CVs, Centers' progress reports), detailed protocols and data abstraction procedures were developed that included specific coding instructions for how materials were to be prepared, entered, and validated. During the period of February 2015 through March 2016, CSI collected and examined 75 biosketches/CVs of the Program's identified key personnel. Data related to awards, honors, service, professional advancements, NIH funding and non-NIH funding were coded and analyzed to support the assessment of evaluation questions. We also used annual progress reports produced by the Centers to describe both the progress made to date and plans for the following year. In the content analysis of these reports, we coded and analyzed descriptions of the accomplishments of the research project, including aims, studies and results, significance, plans, and publications and other project generated resources.

Bibliometrics, Altmetrics, and Publication Analysis

A major source of information accessed, organized, and analyzed were the scientific publications produced by the Centers. The evaluation of scientific work is among the key driving forces behind modern scientific advancements. We queried each of the Centers to submit a list of publications from Center personnel, of which could be attributed to the Program, for the period under evaluation (August 2010-August 2015). From approximately 1000 publications, we employed a structured process to manage the bibliographic output, prepare the records for citation searching, and access the Web of Science (WoS) and other sources to capture the accrued citation patterns of the set.

Using generally accepted bibliometric analytical methods based on computation of normalized indicators, we produced a number of visualizations and metrics to document the changes and impact of the scientific corpus produced by the Centers. This analysis focused on a number of variables to permit the assessment of the degree of contribution of the Program's scientific output:

- **Distribution scientific output**: the spread of peer review publications across journals, fields, topics, and geographic areas.
- **Position of scientific output**: the relative location of Program publications in the hierarchy of journals within fields and science topics.
- **Collaborations**: the co-author relationships across different fields and science topics of relevance to the Program.
- **Structure of scientific output**: the related areas of scientific investigation at Program and their connections to other areas of which the both influence and are influenced.
- **Status of scientific output**: the extent to which Program publications are recognized by scientists in fields in which they are located.

Altmetrics were also collected on the scientific publications submitted by the Centers as part a comprehensive assessment of scholarly impact. Altmetrics comprise metrics based on the integration of social media tools that can inform broader, faster measures of impact, and can serve to complement traditional citation metrics. Altmetrics were used to describe how the research output the Centers was being publicly shared via social media platforms. Several "event categories" were identified as

appropriate for the research outputs of the Centers, with events defined as specific actions applied to articles, such as bookmarks or tweets. Events were collected of a certain type and event counts were simply the total number of events of that type to date. Our analysis produced distributions, categorization typologies, and cross-event patterns.

A third source of information linked to the scientific output compiled from the Centers were authors listed in the biographic records of the submitted papers. This list included key personnel from across the Centers as well as other researchers who contributed to the publications. We used this list of co-authors as a source of data to analyze collaborations. Analysis of co-authorship was determined to be a reasonable proxy for documenting collaborations internally and externally to the Program. We used network analysis methods to examine the changing patterns of collaborations over time. Both visualizations and metrics were generated to assess the growth and development of collaborative work.

Web Analytics/Dissemination Activities

We used web analytic data to document and describe the significant element of dissemination through external access to web resources. Because we did not have direct access to the primary source of web analytic data, we had to rely on the Administrative Coordinating Center at UCLA to supply yearly summaries of unique visitors, web-pages accessed, average pages per visit, and the country of origin. These data were compiled and analyzed relative to the question of how the Program was sharing resources virtually. In addition, the Administrative Coordinating Center at UCLA provided us with information related to the location of scientific meetings throughout the world as well as major scientific meetings/workshops sponsored by NHLBI Proteomics Centers Program.

Quality Assurance Procedures

The evaluation team initiated strict quality controls to ensure that that evaluation meets standards of accuracy, feasibility, propriety, and utility. Methodologically, formal quality control procedures attended to ensuing data integrity, security, and confidentiality through all phases of the evaluation project. The quality controls were organized and managed by the evaluation team in the following five areas: 1) internal and external communication; 2) guidance documents (such as standards of practice); 3) internal documentation; 4) audit protocols; and 5) computer security.

1. **Communication:** CSI maintained consistent communication with the NHLBI Project CORs regarding progress on a bi-weekly basis and ensured the CORs were alerted to any problems encountered as soon as possible. Internal CSI evaluation team meetings occurred weekly to review project tracking, deliverable status, and conduct problem-solving as necessary.

2. **Documentation**: Clear, written expectations for the collection, processing, maintenance, storage, and delivery of data was established and maintained internally. CSI clearly articulated and documented the utilization of systematic procedures and processes for initial quantitative data abstraction and qualitative data collection, as well as verification and validation of data compilation and entry. Routine written updates regarding the application of quality control procedures were forwarded to the CORs on a quarterly basis.

3. **Guidance Documents**: Detailed standards of practice (SOPs) and associated guidance documents were developed to direct the quantitative data abstraction and the qualitative data reduction procedures. The guidance documents for each of the data collection tasks included specific instructions for the data compilation, entry, and management processes. As a verification step, data summaries collection summaries were prepared and included as part of the quarterly reports for verification of progress. Any deviation from the SOPs and guidance documents triggered an internal process for revising and editing the databases accordingly. The internal process included

an internal evaluation team meeting to review and edit the SOP. All revisions and edits were reported to the CORs as necessary.

4. **Audit Protocols**: Audits for the interview data were implemented in two phases. First, the audits followed the completion of a specific checklist to review the interview output. In order to gauge the comprehensiveness of the interview in relation to the guide, a detailed checklist indicating presence (or absence) of specific contextual information and content was completed with every third interview during the initial stages of the process (first 30 interviews completed), for each respective interviewer, separately.

   Interview output that did not meet specific elements on the checklist were flagged and submitted for review by the internal evaluation team for follow up with and correct information or fill in missing data. The review process enabled the interviewers to make adjustments during the early stages of the process of interviewing. Therefore, completion of the checklist was done at wider intervals as the team proceeded with interviews, specifically every 6th interview. The second set of audits for the qualitative interview data was conducted by the Senior Evaluator to ensure the quality of transcription by comparing the transcribed interviews against the actual recording. A random sample of interviews were selected from the final set (n = 8) of transcriptions and were reviewed for accuracy. Findings were reported and reviewed during internal weekly evaluation meetings, which enabled the team to catch and correct problems as necessary.

   Audit procedures for the quantitative data included the Senior Evaluator's review of the compilation procedures to ensure consistency in the populated quantitative database with the expectations outlined in the SOPs. A missing data audit was periodically conducted by the evaluation team to identify any issues regarding completeness and accuracy of data elements. Missing data was reported and reconciled during weekly internal evaluation team meetings.

5. **Computer Security Protocols**: Computer Security included use of an internal local area network (LAN) and use of the Internet over a dedicated T1 line, which includes multiple internal back-up processes on its data systems, and several servers to enable dedicated technology. CSI used offsite secured data storage as well, to assure data security.

Data Limitations

There were some limitations in both the qualitative and quantitative data that should be considered when reviewing the evaluation findings and conclusions. Data limitations were observed in three primary areas: 1) access and availability, 2) quality and gaps, and 3) precision of data.

1. **Access and Availability**: Data were limited in regard to accessibility of the web-analytics, and the biosketch/CV collections. In particular, data pertaining to Google analytics for each individual tag, such as tools and links by year, and summaries of the distribution of program products at conferences were unavailable. Therefore, caution should be taken into account when considering the findings pertaining to the dissemination patterns of the Program. Additionally, *biosketch/CV* documents were inaccessible for 7 out of the 75 identified key personnel of the Program. In some instances, CVs were provided as a proxy for biosketches (14 individuals).

2. **Quality and Gaps**: There were observed gaps in the data specific to the Centers final annual reports, and biosketches/CVs. In particular, Centers' collective annual reports were unattainable by each program year. However, the majority of the Centers' final program reports represented cumulative progress over the course of the entire contract period (5 out of 7 Centers' reports). There were also some inconsistences observed in terms of the reporting structure represented in

the annual reports. For two of the Centers, the final reports indicted progress specific to the final program performance period rather than cumulative reporting. However, annual reports by each program year were complete and available for both Centers, which were analyzed accordingly. In regard to the biosketches/CVs, data related to 2015, were limited due to the data collection time period in that the document collection was initiated in early 2015. Furthermore, the data contained on the biosketches/CVs are dependent upon the accuracy and completeness of those completing the profiles.

3. **Precision of Data**: In terms of the precision of data, limitations surfaced in regard to the following data sources: the external expert sample of the qualitative interviews, altmetrics, and bibliometric data. In particular, the external experts sample revealed a discrepancy in respondents' level of familiarity with the Program. While there are accounts of strong program familiarity among some, others articulated a general lack of knowledge pertaining to the specifics of the individual Centers, and the major advancements of the Program. Therefore, the perspectives of the external experts should be taken into context in regard to their overall frame of reference. In terms of the altmetrics data, there is a level of uncertainly in regard to the degree of value associated with the number of mentions across online platforms in that "Altmetric Scores" are not normalized. However, the percentile scores relative to the mentions captured in the database were produced. Finally, in terms of the bibliometric publication data, there is some ambiguity regarding the explicit expectations of the Program in terms of scientific publication output in that there was not a standardized set of papers to include as a comparative set.

V. Evaluation Findings by Question

This section describes the extent to which the NHLBI Proteomics Centers Program goals and objectives were reached, as well as challenges, lessons learned, and recommendations resulting from implementing the Program. Findings are presented by each of the seven evaluation questions. This section begins with a brief overview of the emergent story of the Centers described by Center Investigators, Trainees and affiliated NIH staff followed by interview quotations organized by theme.

Background on the Development of the Centers and Program

The Program was formed around the creation of multiple integrated Centers, each of which was tasked with forming interdisciplinary teams of investigators and trainees to address a suite of research questions. It was expected that these teams, in turn, would support the development of a shared database and foster an increase in a shared understanding of shared tools and knowledge. Respondents described how the Centers were established by bringing together experts from multiple fields and disciplines, whose efforts were then coordinated in the pursuit of common scientific aims. As the Centers established the interactive model of team research, researchers from a number of domestic and international institutions were engaged as key investigators, subsequently forming global proteomics-focused infrastructure. **Table 3 below** provides a list of the "partnering" institutions, from which key personnel in the Program were connected.

Table 3. List of Participating Institutions

Beth Israel Deaconess Medical Center (Boston, MA)	Massachusetts General Hospital (Boston, MA)	University of Texas (Galveston, TX)
Boston University (Boston, MA)	Royal Institute of Technology (Stockholm, Sweden)	University of Texas Health Science Center at Houston (Houston, TX)
Brigham Women's Hospital (Boston, MA)	Scripps Research Institute (La Jolla, CA)	University of Texas Health Science Center at San Antonio (San Antonio, TX)
Broad Institute of MIT and Harvard (Cambridge, MA)	Stanford University (Stanford, CA)	University of Texas Southwestern Medical Center (Dallas, TX)
Colorado State University (Fort Collins, CO)	Texas Tech University Health Sciences Center (Lubbock, TX)	University of Washington (Seattle, WA)
East Carolina University (Greenville, NC)	University of California at Los Angeles (Los Angeles, CA)	University of Zurich (Zurich, Switzerland)
European Bioinformatics Institute (Hinxton, United Kingdom)	University Family Health Center (San Antonio, TX)	Veteran Affairs Palo Alto Health Care System (Palo Alto, CA)
Florida Atlantic University (Jupiter, FL)	University of Alabama at Birmingham (Birmingham, AL)	Yale University (New Haven, CT)
Harvard Medical School (Cambridge, MA)	University of California, San Francisco (San Francisco, CA)	Zhejiang University (Hangzhou, China)
Johns Hopkins University (Baltimore, MD)	University of Kansas (Lawrence, KS)	
Loyola University Chicago (Chicago, IL)	University of Mississippi Medical Center (Jackson, MS)	

This extensive network was seen as a benefit of the contract as it allowed, as one respondent stated, investigators within each Center "*…to go into entirely new areas and to work on different technologies that complement one another but don't step on each other's toes*" (*Investigator*). Other interviewees commented:

> *… proteomics is an expensive venture because of the instrumentation, techniques, and the technical expertise. And so having a funding mechanism to get that going, so that it can be*

sustained for a long time without having to worry about the ebb and flow of R01s, might support this…and I think at this point it would be great to continue having the support. -Investigator

…they look like groups working very closely together with different kinds of expertise where by pooling our efforts we can accomplish things which one can't do individually. Some of them, many of them within the [Institution], because that's really centered on [Center], but not completely, are within our medical school. They involve both basic researchers and physicians who have patients of different types. We assign people from our laboratories to work very closely together so that they almost become members of both groups. -Investigator

We coordinate well amongst the appropriate interactions that were meant to happen…the purpose of the Center was not to fund four independent labs. It was to fund four labs that work well together such that the coordination occurs more, on what I would call an ad hoc basis, to achieve the known goals. -Investigator

Once the initial building blocks were in place, the Centers employed various methods of facilitating the integration of different investigators and labs within the Centers. Rather than being composed of several "siloed" labs, as mentioned above, the task of the Centers was to create a space, as one respondent stated, "*…for researchers from multiple disciplines to come and play in the same sandbox*"(*Investigator)*. Other interview respondents commented:

The communication between the Principal Investigators is always quite open. We have had monthly meetings and each group will present their work…and some of the developments or some good parts that can be shared by others. –Investigator

Most other grants do a narrow kind of bandwidth in terms of the types of investigators. This really had quite a broad spectrum with people that normally don't talk to each other. We normally wouldn't be communicating regularly with people in the proteomics lab for instance… so that's different…and I just haven't been in a contract in the past that had this broad interdisciplinary nature. –Investigator

In the development of the Program, one of the stated goals was to develop and expand proteomics databases that could be used for proteomics research both within and beyond the Program. Respondents described how data generated at the Centers were organized into centralized databases for use in research. Beyond the individual Centers, respondents also described efforts to create databases at the program-level, including the Cardiac Organellular Protein Atlas Knowledgebase (COPaKB) database, and ways in which the Program drew on and contributed to proteomics databases that were generated elsewhere in the field.

There are collaborations we used to get the right samples and design part of the experiment to get the correct samples for the assays. Then at the end when we actually generate the data, we sit down and discuss the results together and come up with conclusions or a path forward…and then we did it again…sit down and discuss the design, get samples, then later on discuss the results and then plans together. -Investigator

One of the people in our group published a paper in Science where they were basically looking at proteomics in a huge amount of different tissues…looking at all the different proteins that were expressed and how they were expressed. It's more like a reference, and that was a huge amount of work. It was a huge amount of effort, and it really provides a great kind of database for people doing proteomics to use and as a reference. -Trainee

The [Center] has a really a marvelous track record of data sharing, really establishing portals, scientific portals for people to query…and in particular our proteomics group has established tools that other people can apply to their science as well… I think that's a major contribution. -Investigator

At the same time that respondents described the value of databases generated through the Program, in contrast, some expressed concerns about the volume of data being generated. Some respondents mentioned this as a challenge at the center-level, while others described it as a broader issue facing the field of proteomics.

There's so much knowledge. We have so much information from these proteomics studies that we can't possibly work on everything. –Trainee

I'm a little concerned about the volume of data that you can produce. Then you say, "So what?" …the trick is to get to really informative data, not just a lot of data. –Investigator

The structure of the Centers involved the creation of interdisciplinary teams, through which the work of the Centers would be accomplished. While this approach was viewed as generally productive, some challenges did emerge. For example, some respondents described challenges that emerged when project components progressed more quickly than others, resulting in an overall delay. Respondents provided multiple examples of how interdisciplinary teams were organized and activated within the Centers.

Our [Center]…is run by one of the top ten people in the world for biological mass spectrometry…my role is to develop proteomics technologies and apply them through basic research in our group as well as collaborative research within other groups at [Center] or other universities to study different types of disease models using proteomics to gain a better understanding of the molecular mechanism of disease. – Investigator

I think the most important role that I've felt…at [Center] is to bring really talented top scientists who are used to working on their own stuff into a community without politics to allow us to really work at the team level. I think that was the strength of [Center] and I think that remains a strength of the [Center] group. That was both at the training level where they bonded. We would have our monthly contract meetings and it was full, packed all the time. And there was a lot of interplay between the various groups. –Investigator

Building from the previous comment, there were also many examples of how trainees were integrated into these teams, though some respondents, mostly trainees, described how they tended to work in isolation on their specific tasks or within their labs.

When I came to work at [Center] as a post doc, I worked primarily in [Person's] lab, who was also involved in it...but I wanted to do some more proteomics training and to include proteomics in my research. So I joined [Person's] lab as well, and I had a desk and lab space as well. I commuted back and forth because it was at two different branches. -Trainee

So one thing about this Center and the training that I've received is it makes me less timid about doing clinical research. In grad school, having no experience at all, you're timid and you're like, "clinical, oh my gosh." So now I don't feel that anymore. I feel better about trying to reach out to some clinicians. -Trainee

I know that there was a study that was more clinical or translational-focused, but I don't know what was happening with any of the human side. I was purely on the mouse side. -Trainee

Within the Centers and, to a lesser degree, across the Centers, respondents provided descriptions of how investigators from different fields and backgrounds came to utilize shared tools and operate from shared understanding. While, on the whole, respondents praised the process of learning and integration of knowledge within the interactive model, some challenges were also mentioned. These included issues that arose in the day-to-day interactions of investigators across teams, and issues that arose when components of projects did not progress at the same pace. This was particularly evident in the diverse skills of trainees.

For example, we got together once a month so everybody was there. So we could discuss if there were new things being played out in terms of laboratory work and tissues that could be moved to the mass spec core…we are all using different approaches. We're all playing with different ideas, but some of those could be moved over to the core such as their work with I Track and TNT to better help quantification. -Investigator

…my group actually had an exchange with [Person]…I sent a post doc into [Person] lab to learn fractionation…that's one way. Then of course when we meet for conferences and we have exchanges. Another example, [Person] was invited to our [Center] to give a talk because we know what we expect, we can make these invitations and then [Person] can share her knowledge through that. -Investigator

Because the Program is not only about basic science, it's interdisciplinary gathering…it's a painful transition because for me, originally from a chemist kind of background, an electrochemist background…I always had this black and white mindset. I mean, either you'll get this compound or you're not, either you'll characterize this structure or you're not, but biology is not like that…it's more like a spectrum. So I think all the people I'm involved with, they really brought me to today, they helped me to remold my mindset. -Trainee

There were, however, a smaller number of respondents who saw the sharing of tools and knowledge as a challenge.

I wished I had a little bit more access to instruments; that was difficult. They are always worried that we will mess up their instruments with our samples. I mean to some degree, we do that, because it's a biological sample so it's not as clean… chemically clean as a standard, but that comes with the nature. It would have been good to have a little bit more basic support from the instrument people. That interaction was a little rough. -Investigator

Within the cancer Centers… another consortium project, each of the Centers are developing targeted mass spectrometry-based assays for various proteins of interest. And there's a strong effort to make sure that there isn't unnecessary duplication of efforts in terms of multiple Centers creating assays to exactly the same protein and also deposition of the, first off, letting the other Centers know what they're developing assays for, and then depositing the results of those configured assays in a portal. We actually developed and set up an assay portal that is on the [Institution] website for the [program]. That didn't happen within the heart, lung, and blood program. There were other Centers that were using targeted mass spectrometry to develop assays, targeted assays, but those efforts happened very independently, and there wasn't an attempt to try to coordinate that in any way or to capture the information… That, to me, is a failure of the Program. -Investigator

Furthermore, at the program-level, some respondents described the work of their Centers as being unique or removed from that of other Centers. This was perceived as a barrier to the sharing of knowledge and tools across Centers.

> *I feel like yes, we always hear each other's work at these meetings, but I feel like the Centers are pretty different in what they do…and so and I feel we do something very different from other Centers. It's not like we don't use what they developed…it's just that, what we do is very different from what they do basically. –Investigator*

> *I think one of the things with the consortium that didn't work, and that is, that for 12 years we were involved with the consortium, we were the only non-mass spec-oriented Center. I don't feel like our work was particularly heavily influenced by any of the other Centers at all. We largely went to the meetings, we listened to what people presented. Sometimes it was really cool, but it just didn't fit very well to what we were doing, and so I don't feel like we benefited a lot from them or vice versa. –Investigator*

> *…our Center was the only one that was interested in respiratory disease. The other Centers were all heart or pulmonary hypertension, and so we were just in a different sort of area. We thought about some of the technologies that other Centers had developed and tried to apply them to our problem, but they were never a mainstream component of our research programs or our technologies. –Investigator*

After providing some feedback from stakeholders about the background and context to the development of the Centers and Program, each of the seven evaluation questions will be further examined throughout this section.

The Extent the Program Developed and Shared New Tools and Technologies

The question of the extent to which the Program developed and shared new tools and technologies was assessed in two parts. First, we sought to understand the Program's ***development*** of new tools and technologies, focused on both creation and improvement. Our inquiry into the development of new tools and technologies was facilitated through key informant interviews and a review of documents to establish examples of highly recognized results. Second, we sought to understand the ***sharing*** of these new tools and technologies by researchers through web-based platforms as well as in face-to-face exchanges. For this, our inquiry was supported by key informant interview data, web-analytics, altmetrics, and document review.

The Centers engaged in multiple efforts to expand the use and appreciation of proteomics by the broader scientific community. This involved, primarily, influencing investigators from other fields to incorporate proteomics into their work, and the development of tools and methods designed to have a broad appeal. Respondents described several new tools and technologies developed as part of the Centers' scientific agenda. These ranged from new protocols or workflows developed for very specific tasks, to new tools that were applied in animal and/or human models. At the program-level, respondents described how the Program allowed new tools to be vetted and improved. Respondents also described how existing tools were improved through the Program, resulting in faster, more accurate, and higher performing tools. When queried about the significant tools and technologies produced by the Program, respondents highlighted several recognizable advancements:

> *Our Center developed this MRM-based quantitative approach. That is operation of multiple reaction monitoring. So this is a very unique quantitative approach. And if you have internal*

standards, which is also isotope label internal standard, you will be able to absolute quantify how much phosphorylated proteins you have in your samples…So there's a strong expectation that the kind of technologies, the assays, and so forth are progressing towards the clinic. -Trainee

I think the application of CyTOF technology is huge…cumulative evidence suggests that there is immune dysregulation [in pulmonary hypertension] and this technology has really allowed us to get to the bottom of that…and it's huge now. People are applying CyTOF all over the world… I think this is the singular most… major discovery. - Investigator

…from what I understand, the biggest advancement to come out of our lab, and I think it would definitely be within the top three that came out of the entire Proteomics Center during the past five years was the development of this protocol where we tag the proteins with different mass tags…essentially, in order to be able to quantify – we're looking at oxidized cysteines. -Trainee

Another area we spent quite a bit of effort on is developing top-down proteomics now to be able to look at the intact proteins and cross-linking to be able to see what the spatial relationship is between different components and complexes… But I think these kinds of things, then we're looking also at developing things that can become clinically useful assays. We've developed a top-down approach for looking at hemoglobins and both for sequence variance and modifications where we can now take very small blood samples and use MALDI time-of-flight mass spectrometry to get a very fast analysis of hemoglobin variations that we hope will help to get this into the clinic. -Investigator

Investigators and trainees described their participation in conferences and public lectures, through which they highlighted the work of the Program as a way to share new tools and technologies. Some external respondents (Outside Experts or representatives from NIH/NHLBI) also mentioned notable presentations given by program investigators. Beyond presentations, some Centers also engaged in activities designed to make the work of the Program more visible to the public, including the production of podcasts, videos, and newsletters.

Respondents also described how data generated at the Centers were organized into centralized databases for use in research. The sharing of data and results was an important part of the Centers' work and a programmatic expectation at the center- and program-level. Generally, the Centers engaged in efforts to increase the use and application of proteomics. Some Centers offered trainings for scientists outside of the field to learn how to use their tools and methods, while others relied on websites, conferences, and publications as methods of engaging the scientific community. Within the Program, new research tools were shared across the Centers and between individual investigators, through both formal (i.e., websites, newsletters, presentations at PI meetings) and informal (i.e., discussions and personal connections) means.

Respondents also described how tools were improved through this process of dissemination and sharing. Centers within the Program also prioritized the dissemination of new research tools to the field, through various means and to various degrees. Publications were a common method of dissemination. Additionally, Centers disseminated tools through active outreach (including meetings and symposia) and by making tools and databases publically available online. Beyond the individual Centers, respondents also described efforts to create databases at the program-level, including the COPaKB database. At the same time that respondents described the value of databases generated through the Program, some expressed concerns about the volume of data being generated. Some respondents mentioned this as a challenge at the center-level, which others described it as a broader issue facing the field of proteomics.

While the Program played a role in creating new databases, it also drew on and contributed to proteomics databases that were generated elsewhere in the field.

As a significant resource for the field of proteomics, the COPaKB database was created by the Program as a unique resource to facilitate the discovery of novel biological insights from proteomic datasets. The COPaKB database is a consolidated cardiac proteome knowledgebase with novel bioinformatics pipeline and web portals, thereby serving as a new resource to advance cardiovascular biology and medicine serving as a centralized platform of high-quality cardiac proteomic data, bioinformatics tools, and relevant cardiovascular phenotypes. Currently, the COPaKB database features 8 organellar modules, comprising 4,203 LC-MS/MS experiments from human, mouse, drosophila, and Caenorhabditis elegans, as well as expression images of 10,924 proteins in human myocardium. In addition, the Java-coded bioinformatics tools provided by the COPaKB database enable cardiovascular investigators in all disciplines to retrieve and analyze pertinent organellar protein properties of interest. It was intended to provide an array of intuitive tools in a unified web server, so that nonproteomics investigators can conveniently collaborate with proteomics specialists to dissect the molecular signatures of cardiovascular phenotypes.[10]

The web-based COPaKB portal went live in the Program's second year (May 2011). Using web analytic data, we assessed the frequency in which others accessed the information contained on the site over a four-year period. As shown in **Figure 2 below**, COPaKB access increased steadily every year, both in terms of the number of unique visitors and the number of web pages accessed on the site. Indeed, from the end of Year 1 of the web analytic data tracking to the end of year, a 409% increase was observed in the number of visitors and a 281% increase in the number of page views. The largest increase was noted between Year 1 and Year 2, where the number for both visitors and page views about doubled. Furthermore, users from 116 countries were identified as accessing the COPaKB database over the four-year period.

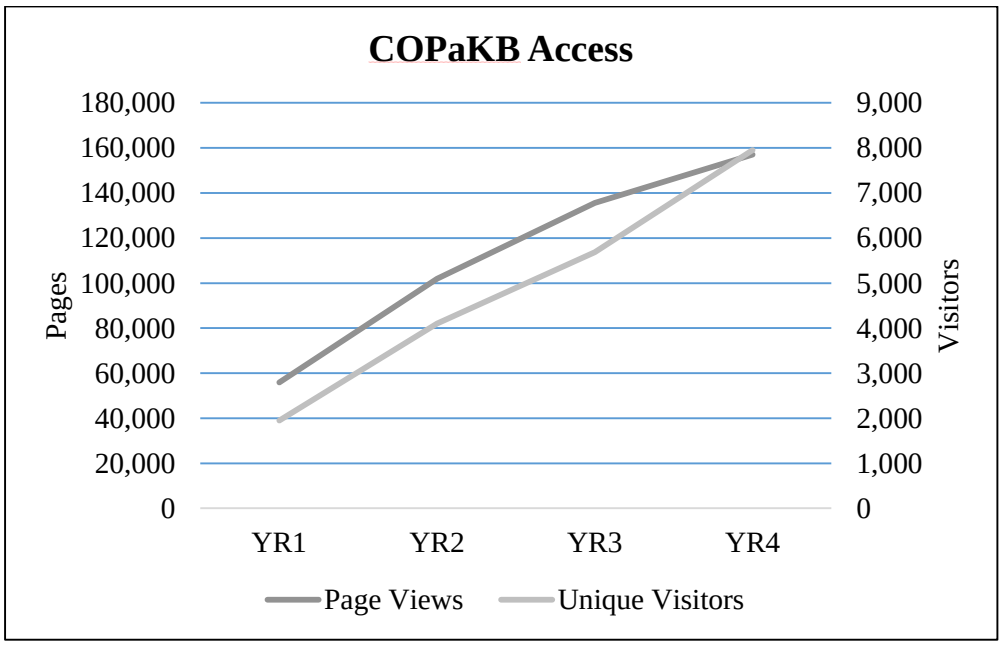


Figure 2. The COPaKB Database Access Trends

[10] Zong NC, Li H, Lam MP et al. (2013). Integration of cardiac proteome biology and medicine by a specialized knowledgebase. *Circ Res*. Oct 12; 113(9):1043-53.

The Administrative Coordinating Center developed a NHLBI Proteomics Center Website (http://www.nhlbi-proteomics.org/). It was officially launched on August 15th, 2011. The website enables the individual Centers to highlight their research achievements and enhance the Centers' visibility within the general scientific community. So far, this website has received enormous responses from the scientific community. Over 80,588 hits have been recorded by Google Analytics since its foundation. The website has been viewed by individuals in 104 different countries and 2,088 cities worldwide.

The presence of the NHLBI Proteomics Center Program scientific output in the hierarchy of peer-review journals speaks to the visibility and credibility of the published scientific work. Assuming the scientists strive to published their work in the most reputable and visible journals in their respective fields, papers residing in the upper tiers of the journal hierarchy signal both quality and attention, as these journals typically have large readership and rigorous peer-review standards.

As another way to understand how the scientific output of the Program was shared across the proteomics research community is through the primary vehicle for communicating scientific advancements. In order to assess these advancements, we examined the journals in which network papers were published. The 975 papers were published in 285 different journals, with the number of publications ranging from 45 to 1. Using the Journal Citation Reports (JCR) Impact Factor as means for organizing the journal hierarchy and evaluating the position and presence of the Centers' publications, we examined the top tier of journals for which Center papers were found. In **Table 4 below**, a list of journals with three or more publications form the Centers indicates the distribution of highly cited papers (i.e., top 1%) across the journal hierarchy is shown. **Table 4 below** also specifies those journals ranked in the top 10% of all scientific journals indexed by the Institute of Science Index (ISI), based on the 2014 JCR Impact Factor Rank (those above the dotted line in the table).

A total of 57 of the 285 journals publishing Center authored papers were found at this ***very high level***. The total number of papers published in the top 10% of all journals was 518 or 53.2% of all 975 papers. Thus, more than half of the scientific output of the NHLBI Proteomic Centers Program appeared in the upper 10% of all indexed scientific journals. A review of the top journals listed in the table suggest the scientific output is being disseminated across a broad scientific community, as journals in this list include *Nature* (estimated readership 424,000 people) and *Science* (estimated readership 570,400 people), are widely circulated.

The 518 papers received 13,607 or 66.7% of all citations attributed to Center publications. In addition, 36 or 97.3% of the top 1% of highly cited papers published by center researchers were found in this group of 57 journals. The average JCR Impact Factor of this select group of journals was 11.40. Based on the Journal Normalized Citation Score (JNCS) it is possible to determine how the set of papers performed relative to the overall publications in that specific journal. For example, the set of Center papers published in *Nature Biotechnology* was cited 500% above what was expected for papers published in that journal (adjusted by year). On balance, the scientific publications of the NHLBI Proteomic Centers Program were visible and accumulating attention from other researchers via citations at the upper most level of the journal hierarchy when considering all journals published worldwide.

Table 4. Distribution of Highly Cited Papers (Rank-Ordered by JCR Impact Factor)

Journal	P	C	MCS	JNCS	# in Top 1%	IF	JCR IF Rank
NAT BIOTECHNOL*	**9**	**1375**	**152.78**	**6.08**	**6**	**41.514**	**6**
NATURE*	**5**	**356**	**71.20**	**1.85**	**3**	**41.456**	**7**
NAT REV GENET	4	241	60.25	2.02	0	36.978	11

SCIENCE*	**5**	**529**	**105.8**	**1.02**	**1**	**33.611**	**16**
CELL*	**14**	**973**	**69.50**	**0.96**	**4**	**32.242**	**20**
NAT METHODS*	**7**	**281**	**40.14**	**0.91**	**5**	**32.072**	**21**
NAT MED	7	962	137.43	2.02	3	28.223	27
CELL STEM CELL	4	39	9.75	0.82	2	22.268	37
NAT IMMUNOL	3	664	221.33	2.66	0	20.004	51
CELL METAB	7	405	57.86	1.61	0	17.565	64
J AM COLL CARDIOL	9	88	9.78	5.46	0	16.503	77
SCI TRANSL MED	10	146	14.60	0.86	0	15.843	82
CIRCULATION	14	349	24.93	0.55	0	15.073	106
J CLIN INVEST	12	369	30.75	0.93	0	13.262	124
AM J RESP CRIT CARE	8	157	19.63	0.73	0	12.996	126
J AM CHEM SOC	3	67	22.33	0.81	0	12.113	143
NAT COMMUN*	**6**	**112**	**18.67**	**1.26**	**1**	**11.47**	**156**
CIRC RES	34	760	22.35	0.90	0	11.019	164
MOL SYST BIOL	4	198	49.50	1.82	0	10.872	166
BLOOD*	**7**	**204**	**29.14**	**0.94**	**1**	**10.452**	**173**
NAT REV RHEUMATOL	4	59	14.75	1.69	0	9.845	194
P NATL ACAD SCI USA	17	396	23.29	0.93	0	9.674	198
NUCLEIC ACIDS RES*	**9**	**937**	**104.11**	**6.61**	**3**	**9.112**	**227**
CELL REP	5	72	14.40	1.18	0	8.358	260
CLIN CHEM*	**4**	**128**	**32.00**	**1.52**	**2**	**7.911**	**282**
ARTHRITIS RHEUM-US	5	113	22.60	1.34	0	7.764	291
ANTIOXID REDOX SIGN	12	176	14.67	1.06	0	7.407	316
J HEART LUNG TRANSPL	3	115	38.33	3.38	0	6.65	373
CHEM BIOL	6	41	6.83	0.84	0	6.645	374
MOL CELL PROTEOMICS*	**44**	**568**	**12.91**	**1.48**	**3**	**6.564**	**377**
HYPERTENSION	4	70	17.50	1.83	0	6.499	386
J NEUROSCI	4	53	13.25	1.37	0	6.344	405
SCI SIGNAL	4	148	37.00	1.37	0	6.279	413
ARTERIOSCL THROM VAS	10	77	7.70	0.52	0	6.008	457
CARDIOVASC RES	7	72	10.29	1.01	0	5.94	466
FREE RADICAL BIO MED	12	131	10.92	1.04	0	5.736	497
ANAL CHEM	24	192	8.00	0.94	0	5.636	517
ACS CHEM BIOL	3	7	2.33	1.16	0	5.331	582
J MOL MED	4	10	2.50	1.05	0	5.107	617
FASEB J	4	28	7.00	0.82	0	5.043	638
BIOINFORMATICS	7	58	8.29	0.85	0	4.981	655
J IMMUNOL	9	71	7.89	0.73	0	4.922	677
PHYSIOLOGY	3	35	11.67	0.88	0	4.857	695
J GEN PHYSIOL*	**3**	**78**	**26.00**	**2.57**	**1**	**4.788**	**714**

MOL CELL BIOL	4	156	39.00	1.86	0	4.777	716
J MOL CELL CARDIOL	18	159	8.83	1.48	0	4.655	744
PLOS COMPUT BIOL	3	5	1.67	0.48	0	4.62	752
CIRC-CARDIOVASC GENE	7	87	12.43	0.88	0	4.631	761
AM J PATHOL	5	64	12.80	1.10	0	4.591	767
J BIOL CHEM	28	491	17.54	1.48	0	4.573	772
J VIROL	4	50	12.50	0.67	0	4.439	837
J AM HEART ASSOC	3	36	12.00	1.36	0	4.306	883
J PROTEOME RES*	**45**	**469**	**10.42**	**1.14**	**1**	**4.245**	**904**
AM J PHYSIOL-LUNG C	4	40	10.00	1.43	0	4.08	973
BMC GENOMICS	3	12	4.00	0.36	0	3.986	1023
BIOPHYS J	7	76	10.86	1.83	0	3.972	1031
J PROTEOMICS	8	52	6.50	0.86	0	3.888	1089
AM J PHYSIOL-HEART C	17	163	9.59	1.77	0	3.838	1119
PROTEOMICS	38	329	8.66	1.27	0	3.807	1141
ARTHRITIS RES THER	6	120	20.00	1.72	0	3.753	1183
CLIN IMMUNOL	4	24	6.00	1.90	0	3.672	1237
METHODS	4	37	9.25	1.35	0	3.645	1260
DATABASE-OXFORD	4	19	4.75	3.56	0	3.372	1494
VIROLOGY	3	34	11.33	0.82	0	3.321	1545
PLOS ONE*	**41**	**427**	**10.41**	**4.68**	**1**	**3.234**	**1632**
DNA REPAIR	4	39	9.75	1.57	0	3.111	1745
RESP RES	3	20	6.67	0.51	0	3.093	1761
ACS COMB SCI	6	35	5.83	3.05	0	3.032	1831
BIOCHEMISTRY-US	4	51	12.75	1.19	0	3.015	1853
PROTEOM CLIN APPL	23	81	3.52	1.17	0	2.956	1915
J AM SOC MASS SPECTR	6	34	5.67	0.60	0	2.945	1925
CYTOM PART A	6	101	16.83	2.98	0	2.928	1946
TRENDS CARDIOVAS MED	3	22	7.33	1.36	0	2.906	1970
BMC SYST BIOL	4	55	13.75	1.28	0	2.435	2728
BIOCHEM BIOPH RES CO	4	17	4.25	0.72	0	2.297	2978
J CARDIOVASC PHARM	3	74	24.67	2.98	0	2.135	3297
INT J MASS SPECTROM	4	38	9.50	1.31	0	1.972	3693
CTS-CLIN TRANSL SCI	4	56	14.00	3.23	0	1.43	5318
CURR PROTEOMICS	4	5	1.25	0.56	0	0.635	8974
ARTHRITIS RHEUMATOL	6	17	2.83	0.61	0	n/a	n/a

Note **(*)**: Journals in bold are those for which the top 1% of highly cited network papers are published.
Note: Journals above the dotted line are in the top 10% of all journals indexed by ISI.

As the scholarly usage of blogs, social media platforms, and other online communication channels has become increasingly commonplace, measuring the online attention paid to research and researchers is important. Altmetrics (alternative metrics) are one of the elements in a comprehensive assessment of

scholarly impact, particularly as it pertains to online presence. Although still in the early stages of development and standardization, altmetrics are viewed as a supplement to understand widespread scientific dissemination. Altmetrics comprise metrics based on the integration of social media tools that can inform broader, faster measures of impact, and can serve to complement traditional citation metrics.[11]

Presumably, the overall corpus of scientific output generated by the Centers included specific peer-reviewed publications describing the development, validation, and utilization of tools and technologies produced and refined by the Centers. Thus, our use of altmetrics was meant to identify the pipeline of scholarly blogs, social media platforms, and other online communication channels within which sharing of the critical research output flowed, including the tools and technology produced by the Program.

In accessing altmetric data for the NHLBI Proteomic Centers publications, we used Altmetric Explorer, where approximately 12,000 online mentions of individual scholarly articles are found every day.[12] This web-based platform includes comprehensive access to social media (i.e., Pinterest, Facebook, Twitter, Google+), reference managers (i.e., Mendeley, CiteULike), blogs, and mainstream media outlets. To incorporate altmetrics, we identified several "event categories" that are appropriate for the research outputs of the Centers, with events defined as a specific actions applied to articles, such as bookmarks or tweets. **Figure 3 below** shows the distribution of mentions across major categories.

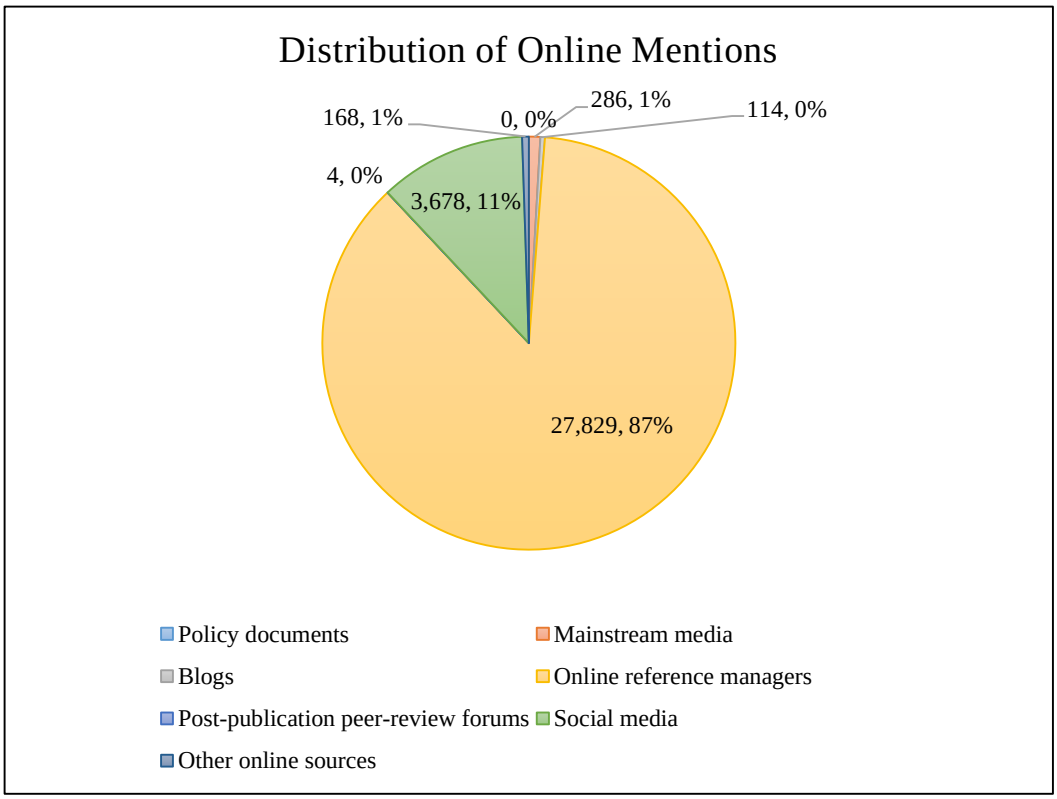


Figure 3. Distributions of Online Mentions Across Major Categories

In general, the NHLBI Proteomic Centers papers published between 2010 and 2015, acquired 32,079 mentions across various online platforms. The majority of mentions were found for reference managers, accounting for 27,829 or 86.8% of the total mentions, followed by social media events, accounting for

[11] Liu J, Adie E (2013). New perspectives on article-level metrics: developing ways to assess research uptake and impact online. *Insights*. 27(2): 153-158.
[12] Altmetric Explorer website (accessed 1 Mar 2016): https://www.altmetric.com/

3,678 or 11.5% of the total mentions. Papers were most frequently mentioned in Mendeley for reference managers, and on Twitter for social media platforms. Mendeley is a reference manager platform that allows for sharing and online collaboration of researchers. As a source of online conversation, Twitter is used by some scholars for scientific sharing and engagement, however non-scholars make up the bulk of users. During the interviews, we found that respondents also commented on the how new tools and technologies were shared in a positive way.

> *We developed an enterprise level top down bioinformatics solution…to sequence proteins and determine sites of post translation modifications to determine specifically the effects of one of our disease models - metabolic syndrome which causes oxidative stress and leads to a cardiovascular stress and cardiovascular phenotype…it's going to be mounted at the super computer facility…which has 10,000 processes cores and it will be free worldwide. And so that's the only one in the world. –Investigator*

> *From a larger standpoint, we also did presentations at national meetings, like the American College of Rheumatology, the American Heart Association, and the European League Against Rheumatism (which is a big European meeting that has 12,000 or 14,000 people), not to mention all the different proteomics meetings that take place…so we were pretty actively involved in those presentations and disseminating information that way. Also certainly from manuscripts as well. –Investigator*

> *We've developed software tools, which are being downloaded regularly by people around not just the country, but the world. I think those are contributions. We've developed protocols and methods, which are being used quite widely. –Investigator*

One respondent did address a critique of the Program and the way that new information was being disseminated:

> *Traditionally NHLBI awards have not had a main focus on dissemination. They just think automatically people will get things published. They see publications as a great dissemination portal, which I agree 100%. What I see as deficiency is if you rely on publication as the only dissemination portal. So dissemination as well as building resources that have long-lasting impact has not been adequately emphasized. Not by this award, not by any other award. And we're going to pay a heavy price for that. –Investigator*

Another, more specific venue for sharing new tool and technologies occurred over the course of the period of inquiry, during three major scientific workshops sponsored by the NHLBI Proteomics Centers Program. Researchers from within and outside of the Program were invited to engage in the advanced technical and scientific topics related to Proteomics. Program researchers from the Centers had the opportunity to share some of their current research and technological developments. The following workshops were held over the five-year period:

- The Omics Integration in Biology and Medicine Workshop in the NIH campus in Bethesda, MD on June 19th - June 20th, 2012. The focus of this meeting was on the emerging field of integrating disparate omic data from genomics, proteomics, glycomics, etc. in order to better understand key biological processes and also improve clinical practice. Discussants focused on identifying the technical and biological barriers in omic integration, with solutions to build a consensus towards data integration in bioscience and to better define phenotypes. Sixty-one (61) internal and external researchers attended this meeting.
- The NHLBI Proteomics Centers Bioinformatics and Database Workshop in Bethesda, MD on August 26th, 2013. The objective of this workshop was: 1) to promote Big Data science and

support data dissemination; 2) to disseminate bioinformatic databases and resources among the seven Centers; and 3) to facilitate the further development of computational tools on proteomics analysis for use by the broader NHLBI scientific community. Fifty-five (55) internal and external researchers attended this meeting.

- The Joint NHLBI Metabolomics/Proteomics Workshop in Baltimore, MD, on August 13th, 2014. These discussions highlighted and promoted cross-disciplinary collaborations and ideas in an effort to move toward a more multi-omics systems biology approach to medicine. Sixty (60) internal and external researchers attended this meeting.

Furthermore, sharing of program information occurred globally, as indicated by the geographic breadth represented in **Figure 4 below**. To facilitate the dissemination goals of the Program, the Administrative Coordinating Center constructed a poster and a brochure to highlight the NHLBI Proteomics Centers. The brochure and poster professionally recognized the individual specialties of the seven Centers. The Administrative Coordinating Center reached out to the global scientific community, covered by more than 12 scientific organizations both nationally and internationally, to showcase this program and its seven Centers, including five councils that belong to the American Heart Association (AHA: BCVS, FGTB, ATVB, 3CPR, and Stroke), four chapters that belong to the Human Proteome Organization (HUPO: International HUPO, US HUPO, Asia Oceania HUPO, and Chinese HUPO), the International Forum of Proteomics (IFP), the American Society of Mass Spectrometry (ASMS), the International Society for Heart Research (ISHR), the Society of Heart and Vascular Metabolism (SHVM), the Proteomics Standards Initiatives (PSI), the British Society of Proteome Research (BSPR), the Scandinavian Physiological Society (SPS), the Society of General Physiologists (SGP), the American Physiological Society (APS) and the American Society for Biochemistry and Molecular Biology (ASBMB).

In addition, the NHLBI Proteomics Centers Program materials were displayed during several Chinese Cardiovascular conferences and Omics conferences. Center posters have drawn attendees' attention and highlighted the scientific aims and milestones of each Center. The Administrative Coordinating Center also distributed the newsletter, *NewsSpot*, which highlighted the latest achievements of the Centers at multiple conferences and scientific organizations around the globe. Moreover, the Administrative Coordinating Center created a master email list containing thousands of researchers' emails collected from the journals (over the past five years) in proteomic and cardiovascular areas. Some of the e-newsletters were distributed among these scientists. On the map in **Figure 4 below** are the 31 locations where 43 scientific meeting were held and the NHLBI Proteomics Centers Program materials (i.e., posters and brochures) were distributed and featured.

Geographic Locations of NHLBI Materials

Figure 4. Locations of Scientific Meeting where NHLBI Proteomic Centers Program Materials Where Featured.

The Extent the Program Created Outputs of Clinical Utility

Based on this extent of exposure of the Program, the outputs of the Program have matured enough to be of clinical utility. Furthermore, the extent of the Program has resulted in inventions, patent applications, spin-off companies, and licensing agreements.

Overall, investigators within the Program reported having made significant progress in the development of tools and generation of knowledge with clinical relevance, and they provided multiple examples of the clinical significance of their accomplishments. Some respondents also described the significance of the work to diseases beyond heart, lung, and blood. At the same time, it was clear from the data that the march toward clinical relevance and clinical applications was a work in progress, and a long road. Within the Program, respondents described the next steps needed to accomplish their ultimate clinical goals. The assessments of respondents who were external to the Program were equivocal; while some were in agreement that the Program had made significant progress on this front, others were much more guarded or skeptical in their assessments.

> *…the implications of the repertoire analysis are that it allows one to look in many different fields, including vaccinology or in any blood disease in which there might be antibodies that are playing a role, and you can look at the repertoire broadly to see what antibodies are being made…That has potential to increase the speed at which we can do studies, I would say in any blood disease, for sure, so anything where blood is involved, and that would include heart disease, stroke, pulmonary disease, certainly cancer, and then in immunology in general. Those antibodies can be developed into diagnostics or therapeutics. -Investigator*

> *you're getting samples from a patient that have been properly collected and processed and cryopreserved, if they need to be cryopreserved, it's a really difficult task, because you have to have a coordinator who's physically in the clinic, and consenting patients, then getting the blood and then processing it quickly and getting it to the right person, and then the amount of infrastructure funding that's needed for that is well beyond what many of these contracts can provide. I feel like we did a lot there for studying the human samples. We probably ran over 100 patients through the different technologies that we developed and have some nice papers that are being written up currently. –Investigator*

Furthermore, respondents provided detailed examples of the process of conducting animal and pre-clinical studies. Within the Centers, respondents described the role of animal and pre-clinical studies as informing or paralleling the clinical work. Relatedly, there were several references to the components of the three-part model all feeding, ultimately, into clinical studies and goals. Respondents from some Centers described specific clinical and translational studies, including studies on human tissues, which had been conducted within their Center. Again, details on the mechanics of this process were provided.

> *Another great example is we just had a nature paper with [Name] on kinase g…the way you activate it is by inhibiting the breakdown of cyclic GMP so cyclic GMP binds the kinase, the kinase does its job to protect the heart…there's a whole bunch of drug trials on inhibiting this PDE which is this thing that stops the breakdown of cyclic GMP…and they're not coming in as well as one thought would be and so… lots of companies are coming to us to say 'Hey. We need to get our drugs on target. Can you help us?' and it's only through proteomics." -Investigator*

> *…there was a very interesting project where the clinicians in the [Institution] provided us with samples of sepsis patients [that died]. So, the proteomics core was able to look at those samples and identify early, early markers that would tell you the progression of the disease and the likelihood of that patient to survive or not. -Trainee*

...I think a couple of major impacts has to do with at first using phosphoproteomic technologies to inform us about heart failure, both new pathways into heart failure, specifically PD9 which [Name] just published but also regulation of the phosphoproteome in heart failure. -Investigator

...we have an ongoing 600-plus patient sample study that we're going to wrap up and I have a total of over 1,000 patients that will have been analyzed by targeted mass spectrometry-based approaches. That's the largest study ever using this technology. I think it will be – I hate to use the word "landmark", but it's certainly this line in the sand saying that you can accomplish clinical-size studies using targeted mass spectrometry. -Investigator

Our first clinical study was a grand success and the publication was out in the summer last year. At this moment we are in the process of organizing another large clinical trial to evaluate additional parameters with respect to both efficiency as well as accuracy of single biopsy samples. -Investigator

...they've identified certain biomarkers that go up or down with heart failure patients. I know now that this particular marker is actually accepted now by the committee as sort of, "Hey, if this particular protein's up it could be a sign of heart failure progression". -NIH

We were able to demonstrate that there are different patterns of proteins found in the airway of patients with severe asthma and that these protein signatures were related to the airway physiology. -Investigator

Basically, we are using the next-generation sequencing technology to better come up with a more detailed picture of immune response in individual patients. And so again, we're doing that in multiple disease types. -Investigator

I think the focus is very ripe for clinical applications…I think it's been largely a program of discovery, but the Program is positioned now to look more at outcomes…And in that, I think it will bear important fruition in terms of saying, "Well, this particular constellation of protein peptides, this particular constellation of metabolites is linked to an adverse outcome". -Investigator

The Extent Research Contributed to Knowledgebase Linkages

The question of the extent to which research in the Program contributed to the creation and integration of a knowledgebase linking changes in proteomes with molecular phenotypes of disease was assessed. There are multiple examples in the data showing how learning developed and evolved across the three core areas (technology, clinical applications, and basic science), and the collective role that these areas played in pursuing the scientific aims of the Centers. Respondents described how, in bringing together experts from multiple fields, the Centers were able to capitalize on their combined expertise in approach projects and questions. At the same time, the integrated model was described as contributing to the development of investigators and experts who were increasingly interdisciplinary in their knowledge and approach. Respondents described how the opportunity to work closely with other investigators from other fields expedited the process of discovery in the Program, as it facilitated communication and exchanges that accelerated the advancement of knowledge and understanding. There was general agreement among respondents that, over the course of the Program, significant growth has occurred in the field. For example, the field has progressed in technical development, engineering research, and clinical

advancement. However, there were differing opinions, particularly by individuals internal and external to the Program, about the extent to which the Program has contributed to these advances.

Investigators at most Centers reported having confirmed linkages between specific proteomes and the molecular phenotypes of the disease or diseases on which they were focused. They also reported demonstrating pre-clinical or clinical uses of new assays or post-translational modifications. Overall, investigators within the Program reported a significant increase in the past five years in the field/area. For example, the Program has been able to influence areas of heart failure research. This combines ways the Program has expedited discovery within the field, as well as how the field has progressed in five years since this iteration of the Program began through examples referenced in the interviews.

> *We've discovered on the order of several hundred novel markers of candidate biomarkers of the myocardial infarction process, markers that have never been pointed to previously, we have developed – so that's number one from a discovery basis, employing state-of-the-art proteomics to discover new markers of myocardial infarction and related cardiovascular diseases. That's a major contribution. -Investigator*
>
> *I think in terms of what we had in mind that the post translational modifications can be very important, we've been able to demonstrate in a number of cases where the overall protein levels have changed only slightly. But significant post translational modifications have changed many fold more. –Investigator*
>
> *There's significant knowledge advances, especially from the perspective that we took, and the creative approach where we're looking at is not the traditional markers due to oxidative stress. We're looking at keeping it in an open plane and using these tools to look up to 18 different post-translational modifications that occur. -Trainee*
>
> *I started collaborating with people that I didn't think I'd collaborate with in terms of like O-GlcNAcylation and things like that, and diabetes, kind of collaborating people like [Name] at [Center]. Because of being able to identify modifications in the mitochondria. –Investigator*

Areas the Program Filled Knowledge Gaps in Heart, Lung, and Blood Disease Biology

The question of the extent to which the Program filled existing knowledge gaps in heart, lung, and blood disease biology was assessed. Respondents described the research goals of their Centers and labs, including goals related to technological advancements, basic science research questions, and clinical goals. While most respondents were able to articulate these goals, some trainees and investigators struggled to state the overall goals of their Center. For some Centers, the plans for research evolved over time. In some cases, this was due to a roadblock that had to be worked around, and in others it was spurred by a new discovery or advancement.

> *So from the technological standpoint, I think that when we started, the ability to measure lots of things and lots of samples [were] poor. I would say that we've contributed very significantly to technologically - how to measure lots of things and lots of samples as quickly as possible. -Trainee*

> *I think the ability to detect novel post-translational modifications that haven't been detected before. It's moved more from qualitative to quantitative now, and that includes being able to look at the ratio of a post-translational modification to an unmodified protein using, for example, SRM or MRM methods, single reaction monitoring. And now that's growing into doing hundreds or thousands of quantifications at the same time. –Investigator*

> *A very important conceptual advance is the first step towards personalized therapy in asthma…so the impact here is that most doctors treat patients with asthma as though it's a singular disease. Our protein analysis studies indicate that is the wrong way to do it and that there are very different physiologies that are present in patients with severe asthma. –Investigator*

> *So the whole thing of working in the Center is that it has substantially shortened the translation time from the maturity of the technology into biology and clinical medicine…the feedback loop of going back and forth has been substantially accelerated…this is where I see things would not have happened so fast if we were working separately. –Investigator*

Presumably, the scientific agenda of the NHLBI Proteomic Centers Program, as articulated by the scientific aims outlined by the Centers, constitute the gaps in the proteomics research base. We assumed that the researchers developing the scientific aims initially were identifying the areas where new knowledge was needed, required, or warranted based on their substantial experience and work in proteomics. In addition, NHLBI as the funder, approved the programmatic scientific agenda. Thus, it is reasonable to conclude that the scientific aims of the Center, and cumulatively the scientific agenda of the Program constituted some of the identified gaps in the proteomics research community understanding of heart, lung, and blood disease biology. Therefore, we extracted the scientific aims form the Centers and organized them by the three areas of integrated research framed by the NHLBI Proteomics Centers Program. **Table 5 below** lists the scientific aims for the Centers, by technology development, mechanistic and molecular studies, and clinical application and represents the overall scientific agenda of the Program.

Table 5. Scientific Agenda of the NHLBI Proteomics Centers Program

Technology Development Aims
Identify novel early biomarkers (~ 80) of myocardial injury in humans as well as animal models through mass spectrometry-based approaches
Develop separation and array-based proteomic technologies to help facilitate protein expression and post-translational studies of asthma, COPD, and respiratory viruses
Develop novel proteomic methods for quantitative post-translational modification analysis and for probing the subproteomes of the cell surface and secretory pathway
Provide new and extended bioinformatic tools and resources centered on human proteins and peptides
Develop extracellular matrix specific proteomic and computational approaches to catalogue extracellular matrix fragment generation and elucidate the degradomic pathways in a mouse model of myocardial infarction
Develop high throughput functional assays to screen extracellular matrix fragments for biological activity in cardiac fibroblasts, macrophages, and endothelial cells in vitro and in vivo
Develop phosphor-flow cytometry and mass spectrometry techniques for the isotope-based analysis of signaling pathways at the single cell level
Develop software for rapid annotation of proteins
Mechanistic/Functional Studies Aims
Clinically characterize the novel early biomarkers of myocardial injury in several small but well phenotyped studies.
Functionally characterize the proteomic changes occurring in myocardial injury using in vitro and in vivo models of ischemic injury.
Interrogate the dynamic nature of the pathways in the mitochondria, including protein composition, protein interaction, post-translational modifications, protein trafficking and protein behavior.
Understand the cardiac proteasomal degradation system, the processing and degradation of targeted substrate repertoire, and the intrinsic principles governing protein quality control in the heart.
Interpret the mitochondrial and proteasomal data on a network scale to ascertain emergent properties that cannot be gained through isolated molecule or single pathway approaches.
Employ genome wide association studies and quantitative proteomic approaches to measure properties of networks and link them to complex traits such as susceptibility to myocardial ischemia and cardiomyopathy
Investigate the pathways and mechanisms of action of airway diseases in lung tissue (mucosal) and blood leukocytes to identify critical causative pathways and possible targets for therapeutic interventions.
Develop diagnostic and prognostic biomarkers for disease severity and treatment sensitivity for allergy/asthma, COPD and RSV infection.
Define the cardiac myocyte proteome and its post-translational modification status in heart failure and diabetic cardiomyopathy
Perform large scale analysis of absolute quantitative data from hundreds of cardiac proteins and their post-translational modifications derived from human samples
Involve the production of a dynamic cell model based on biologic and proteomic data
Identify key extracellular matrix fragments generated after myocardial infarction injury
Determine role of extracellular matrix fragment signaling in LV tissue injury and remodeling
Investigate the role of viral triggers and inflammatory cells in PAH pathogenesis by using rodent models and patient samples
Characterize serum autoantibodies and cytokines to define the autoimmune targets associated with disease initiation and progression, with the ultimate goal of developing “clinically actionable” diagnostics for PAH
Evaluate the contribution of cytokines, chemokines, growth factors, and signaling pathways to PAH in animals and patients using the High-throughput Immunophenotyping using Transcription (HIT) and reverse phase lysate microarray proteomics platform
Develop peptoids as antagonists to perturb cells and as detector molecules to profile lineage specific cell surface molecules and serum autoantibodies
Identify and confirm signatures of putative peptide and oxidative posttranslationally modified peptides in blood and specific tissues in mouse models of heart and vascular disease
Clinical Application Aims
Confirm the presence of the modified proteins identified in the mouse models in relevant and small controlled clinical studies

Prospectively validate the diagnostic, prognostic and decision making utility of the ~20 clinical markers that will emerge from clinical and functional studies
Reveal pathways operational in reverse remodeling in patients receiving LVAD
Identify molecular markers to specify treatment regimens in end stage heart failure
Apply proteomic technologies to three clinical applications, related to asthma, COPD, and respiratory viruses and will include measures of protein/peptide expression arising from multiple oxidative stressors.
Verify diagnostic and prognostic biomarkers for disease severity and treatment sensitivity for allergy/asthma, COPD and RSV infection
Develop plasma markers for risk stratification of patients with heart failure
Generate and disseminate cardiac specific and post-translational modification findings as a resource for human heart failure community
Assess clinical applicability of the identified extracellular matrix fragments to serve as biomarkers of adverse left ventricle remodeling post myocardial infarction
Apply research findings to identify myocardial infarction patients who are at high risk for heart disease/failure
Validate the protein candidates in a larger cohort of individuals derived from an ongoing observational study

The Program's success in addressing these gaps were verified by the respondents. There are multiple examples in the data of how learning developed and evolved across the three core areas (technology, clinical applications, and basic science), and the collective role that these areas played in pursuing the scientific aims of the Centers. Respondents described how, in bringing together experts from multiple fields, the Centers were able to capitalize on their combined expertise in their approach to projects and questions. At the same time, the integrated model was described as contributing to the development of investigators and experts who were increasingly interdisciplinary in their knowledge and approach.

> *…my Center developed the absolute best capabilities currently available for deep proteomics characterization of the plasma proteome. So without that, we wouldn't have been able to discover these markers. So we developed various fractionation methodologies to allow us to dig much, much deeper than previously possible in blood. So this is a major outcome of the Program. –Investigator*

> *...I would say two big advancements are identifying what those extracellular components are that get processed by that enzyme. And then second, not just extracellular, but inflammation is hugely controlled by this enzyme. So how inflammation and fibrosis work together in concert to generate that scar…that's the stuff we're proud of. –Investigator*

In analyzing the extent to which the NHLBI Proteomics Centers Program filled gaps in the existing heart, lung, and blood disease knowledgebase, we modeled the topical network of the scientific output produced by the Centers. To do this, we conducted a network analysis of the author supplied keywords from the publications submitted by the Centers for the period. From the set of 975 papers, the author-supplied keyword network represents a model of the scientific content produced over the period. In total 396 separate keywords were extracted and examined relative to one another as they exist in a relational network.

The density visualization produced by VOSviewer shows the keywords in proximal space found in **Figure 5 below**. Each point on the map has a color that depends on the density of the items at that point. The color ranges between red to blue. The color of the point is closer to red by (1) the larger the number of items in the "neighborhood" of a point and (2) the higher the weights of the neighboring items. Conversely, color of the point is closer to blue by (1) the smaller the number of items in the neighborhood of a point and (2) the lower the weights of the neighboring items. In effect, this serves to produce a "heat map" and color scheme that enables the easy identification of "hotspots" or areas where a high number of items, in this case keywords, have co-occurred.

These hotspots show where a sets of keywords have clustered based on their frequency of appearing in the keyword lists of the same papers. Furthermore, these keywords suggest concentrated activity in the overall research corpus. **Figure 5 below** depicts the density visualization produced for the author-supplied keywords from the NHLBI Proteomics Centers Program scientific output.

Density Visualization of Keyword Analysis

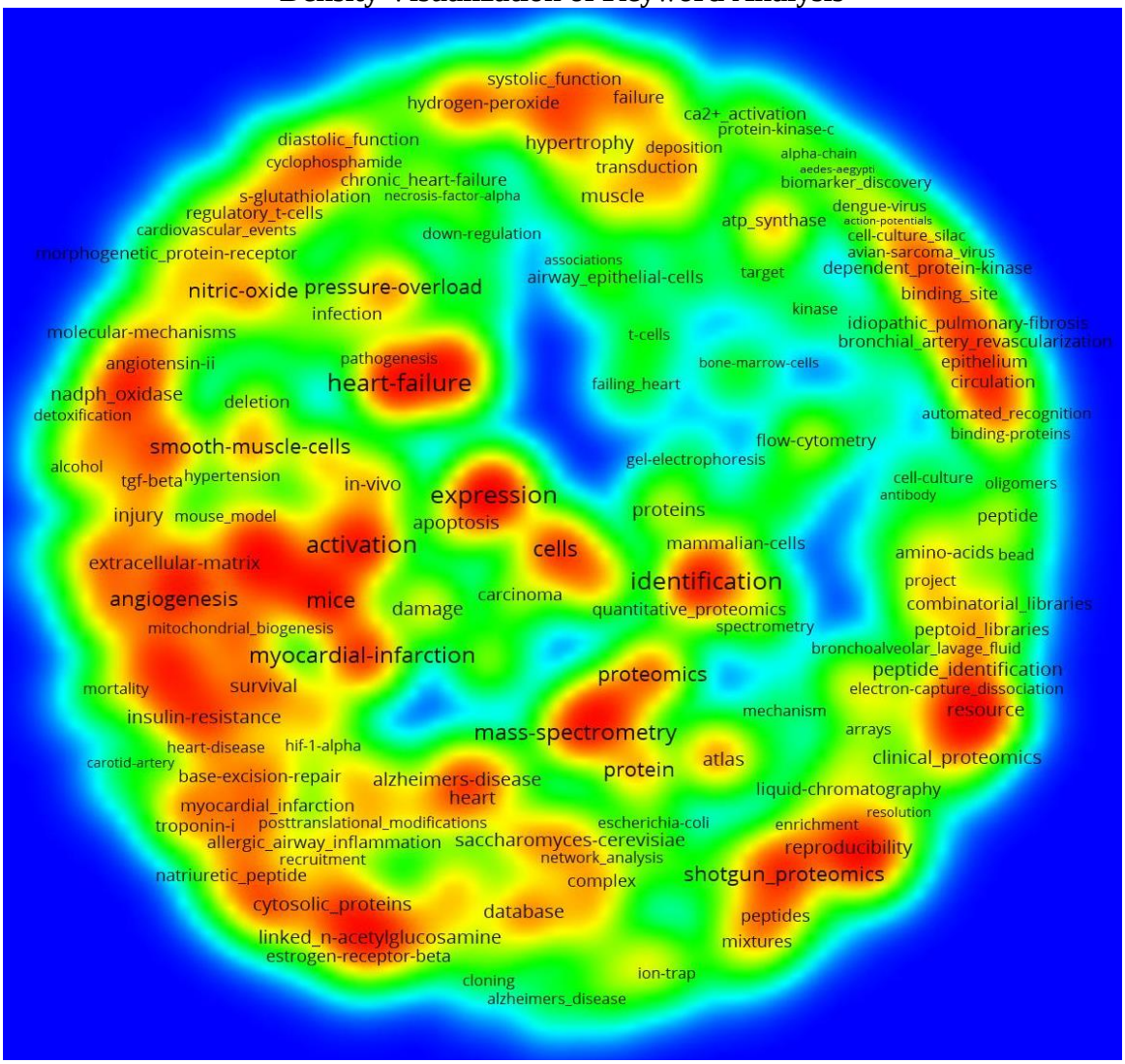


Figure 5. Density Visualization of a Keyword Network from the NHLBI Proteomic Centers Program Scientific Output

The density visualization in **Figure 5 above**, shows the topical areas generally comport with the scientific aims represented in the overall scientific agenda of the NHLBI Proteomics Centers Program. This heat map also highlights groupings of topics across the three focus areas of the Program. We found references related to the mechanistic and functional understanding of the proteome (both interactions and dynamics) included references to the discovery of biomarkers triggered by myocardial ischemia. Several areas included references to protein expression and post-translational modification studies, and a number of specific cardiac proteins and their PTMs (e.g., hif 1 alpha, o linked n acetylglucosamine), as well as the extracellular matrix were linked to other areas in the network. We found references to emphasis on tools

for proteomic research, such as novel mass spectrometry, flow cytometry, and liquid chromatography techniques to plasma analysis. In addition, several areas emphasized methods including identification (i.e., measurement) via quantitative proteomics, network analysis, shotgun proteomics, and peptiod libraries. Finally, we observed several areas emphasizing clinical applicability with topical references to heart failure, injury, cardiac events, angiogenesis, myocardial infarction, and Alzheimer's disease. Specific emphases were identified on airway and ventricle remodeling post myocardial infarction.

On the basis of this plan-to-result review, we can confirm at a high level that the areas of scientific inquiry set forth by the Centers individually and the Program collectively at inception, were manifest in the scientific output produced during the period. The relational network of keywords revealed the complex and integrated coverage of the cumulative scientific work of the Program, and suggest the Center researchers were comprehensive in their efforts to fill the gaps in heart, lung, and blood disease biology initially identified at the outset.

To further evaluate structurally the extent to which the NHLBI Proteomic Center Program addressed scientific gaps in heart, lung, and disease biology we examined the distribution and performance of publications from the Centers within fields of relevance to proteomics research. Scientific papers are published in a range of journals that have varying status in a scientific field. Analyzing the distribution of the paper-journal-field relationship in the scientific literature landscape can reveal much about the prominence of scientific output. This inquiry enabled us to determine which fields the scientific output from the Centers was most prominently featured. We argue in this examination that from any given set, a high proportion of papers in the **top 3 journals** of a field constitutes a highly regarded and influential set of papers for that field. We found the full set of 975 NHLBI Proteomics Centers' papers were distributed across 70 specific fields in total. **Table 6 below** lists the fields with five or more publications and includes the number of Centers' papers linked to the field.

As a means for understanding the level of the Center papers relative to the importance of the journals within each field, we ranked and ordered the journals in each field by eigenfactor score and identified the top 25% (1st quartile) as an indication of the top tier of journals for that respective field. The eigenfactor score is a measure of the total influence of a journal based on cross-citation patterns. The eigenfactor uses the five previous years for the target window and excluded journal self-citations.[13] Thus, the top 25% of journals in a field, based on eigenfactor ranking, is a robust indicator of the prominence of the journal within the field. We also counted the number of papers that were published in the **top 3 ranked** journals in their respective field, further distinguishing the highest level of recognition within each field.

In reviewing these results, some variability in patterns emerge, that allow us to further distinguish the prominence of journals within a field and the number of papers published at the highest levels. For example, in the field of *Hematology*, 54 out of 57 or 94.7% papers from the Centers were published in the 1st quartile of journals for that field, with 50 of the 57 or 87.7% in the **top 3 journals** in the field. This suggested that the set of papers in *Hematology* were present in highly regarded journals. Thus, it could be argued that the collection of papers in *Hematology* was a substantive set of highly regarded and potentially influential set of scientific results.

In contrast, 51 out of 209 or 24.4% of the papers were published in the **top 3 journals** within *Biochemistry & Molecular Biology*. Thus, a much greater proportion of papers in the field of *Hematology* were published in the most prominent journals in the field. These analyses shed light on how well the scientific output of the Centers are positioned in their relative fields as well as how well they are positioned for future visibility and potential influence. Of the list of fields in **Table 6 below**, 10 of the 35 fields had 50% or more of their respective sets published in the **top 3 journals** for their fields. It is this

[13] Bergstrom CT, West JD, Wiseman MA (2008). The eigenfactor metrics. *Journal of Neuroscience*. 28:11433-34.

concentrated set of 238 papers in the 10 fields that have highest level of visibility and potential to be influential in a respective scientific field.

Table 6. Journal Hierarchy by Field (with 5 or more publications)

Field Category	Total Papers	Cites	Cites Per Paper	Papers in 1st Quartile	Percent of Total (1st Q)	Papers in Top 3 Journals	Percent of Total (Top 3)
Biochemistry & Molecular Biology	209	5644	27.00	181	86.60	51	24.40
Biochemical Research Methods	204	2278	11.17	162	79.41	16	7.84
Cardiac Cardiovascular Systems	141	2109	14.96	114	80.85	24	17.02
Cell Biology	113	3953	34.98	95	84.07	23	20.35
Peripheral Vascular Disease	**86**	**1429**	**16.62**	**78**	**90.70**	**46**	**53.49**
Multidisciplinary Sciences	**77**	**1986**	**25.79**	**77**	**100.00**	**62**	**80.52**
Hematology	**57**	**1038**	**18.21**	**54**	**94.74**	**50**	**87.72**
Medicine Research Experimental	48	1715	35.73	40	83.33	21	43.75
Genetics Heredity	41	890	21.71	35	85.37	5	12.20
Endocrinology Metabolism	36	832	23.11	36	100.00	3	8.33
Chemistry Analytical	**35**	**273**	**7.80**	**33**	**94.29**	**24**	**68.57**
Immunology	35	1214	34.69	26	74.29	14	40.00
Physiology	35	410	11.71	28	80.00	17	48.57
Biotechnology & Applied Microbiology	**33**	**1737**	**52.64**	**28**	**84.85**	**17**	**51.52**
Respiratory System	26	534	20.54	23	88.46	10	38.46
Mathematical & Computational Biology	25	204	8.16	18	72.00	12	48.00
Rheumatology	22	350	15.91	16	72.73	6	27.27
Biophysics	**21**	**211**	**10.05**	**16**	**76.19**	**13**	**61.90**
Chemistry Multidisciplinary	18	281	15.61	12	66.67	3	16.67
Oncology	17	468	27.53	12	70.59	3	17.65
Critical Care Medicine	**14**	**272**	**19.43**	**12**	**85.71**	**11**	**78.57**
Pharmacology & Pharmacy	14	186	13.29	6	42.86	0	0.00
Spectroscopy	13	94	7.23	11	84.62	0	0.00
Virology	12	131	10.92	12	100.00	5	41.67
Neurosciences	**10**	**79**	**7.90**	**7**	**70.00**	**5**	**50.00**
Surgery	10	148	14.80	10	100.00	0	0.00
Biology	8	45	5.63	6	75.00	0	0.00
Chemistry Medicinal	8	66	8.25	2	25.00	0	0.00
Chemistry Physical	8	47	5.88	0	0.00	0	0.00
Transplantation	8	155	19.38	7	87.50	3	37.50
Pathology	**7**	**87**	**12.43**	**5**	**71.43**	**5**	**71.43**
Cell & Tissue Engineering	**6**	**47**	**7.83**	**5**	**83.33**	**5**	**83.33**
Chemistry Applied	6	48	8.00	0	0.00	0	0.00
Toxicology	6	58	9.67	6	100.00	0	0.00
Engineering Biomedical	5	23	4.60	4	80.00	1	20.00

Note: Fields in bold had 50% or more of their set in the top 3 journals.

The Extent the Program Investigators Collaborated Within and Outside the Field

In order to examine the extent to which Program investigators have been collaborative within and outside of their fields we used data collection methods. First, we used co-authorship on the scientific output of the Centers as a proxy measure. This enabled us to model the network of collaborators as it changed over the

period of inquiry. Second, we inquired about collaboration during the key informant interviews to provide contextual information and the nature of their collaborative work.

In terms of the patterns of collaboration based on co-authorships, our network analysis revealed the growth of an extensive collaborative network over the period of the Program. As the network accumulated, more researchers engaged in co-authored studies and changes to both the size and cohesiveness of the network was noted. We conducted network analyses to visually represent and assess the collaborative connections of the research output from the Centers. A complete description of the network analysis can be found in **Appendix D.1.8: Co-authorship and Collaboration Analyses**.

The results of the analyses, shown in **Figures 6 through 11 below**, are relational maps of all co-authors listed in the 975 NHLBI Proteomics Centers publications. The different colors indicate clusters of authors strongly connected to each other based on the frequency they appeared together in the author lists of the Proteomic Centers' papers. The colored groups presumably represent "communities of authors" based on the co-author relationships found in the network. The size of the node (i.e., the bubble) and the label (i.e., the author's name) indicated the volume or number of times these authors were listed with others as co-authors. Several of the larger nodes across the maps are well-known principals and investigators, of each Center. This highlights their emergence as central purveyors of collaborative relationships with other proteomic relevant researchers included in the extensive network.

Publication Co-Authorship Network (2010)

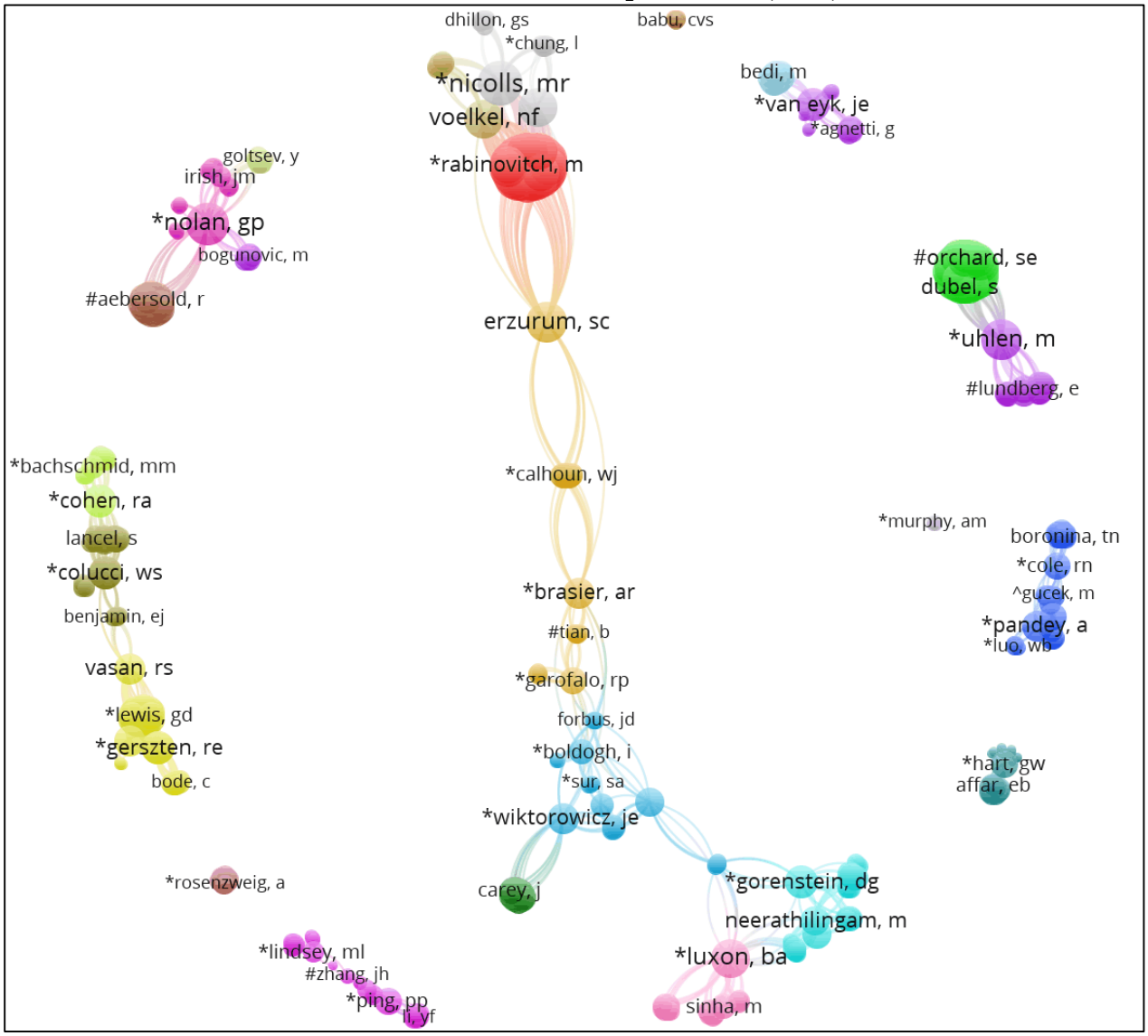


Figure 6. Co-authorship Network of NHLBI Proteomic Centers Publications: All authors 2010.
Note: * = Center personnel; # = Outside Experts; ^ = NIH personnel

Publication Co-Authorship Network (2010-2011)

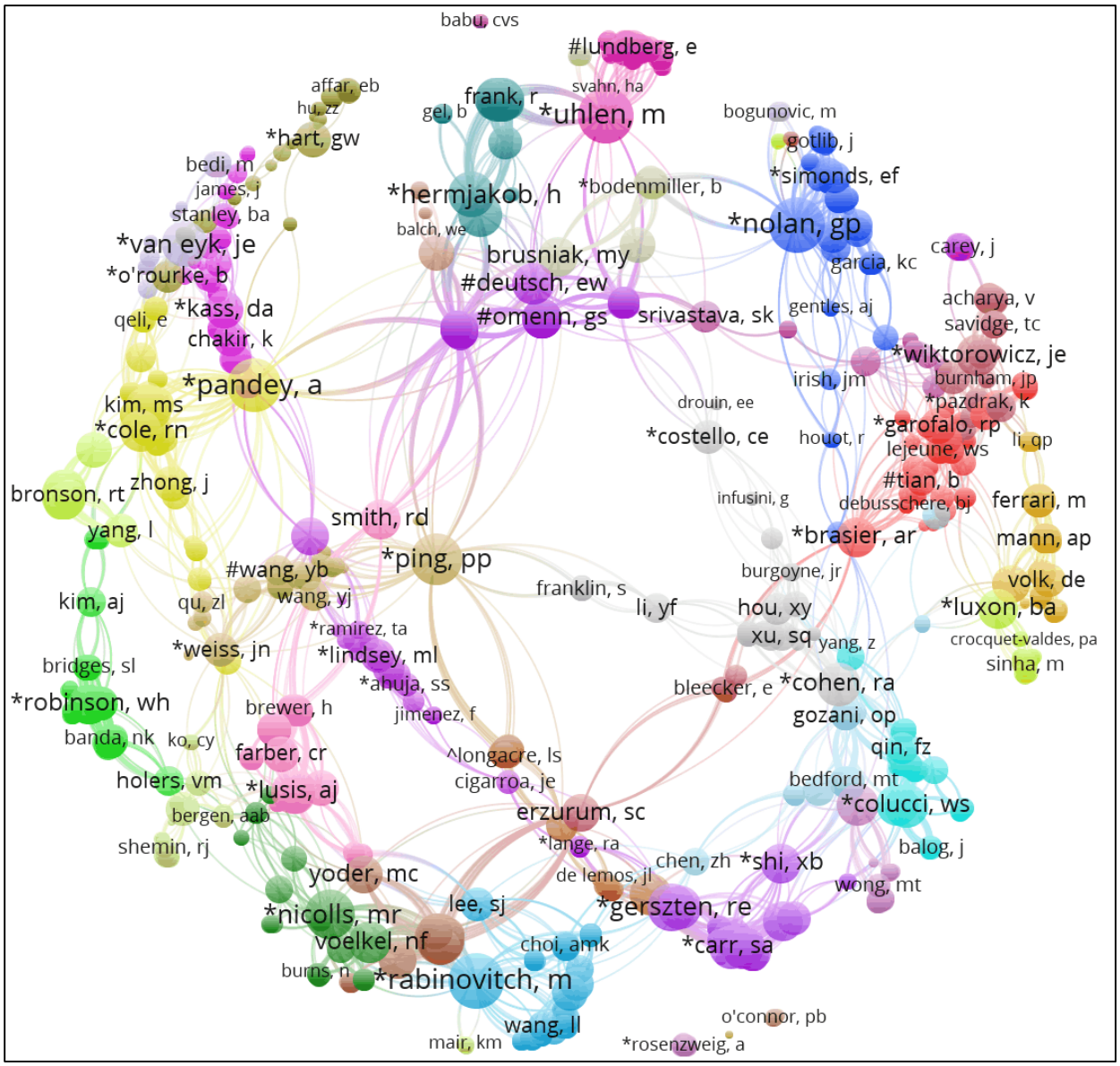


Figure 7. Co-authorship Network of NHLBI Proteomic Centers Publications: All authors 2010-2011.
Note: * = Center personnel; # = Outside Experts; ^ = NIH personnel

Publication Co-Authorship Network (2010-2012)

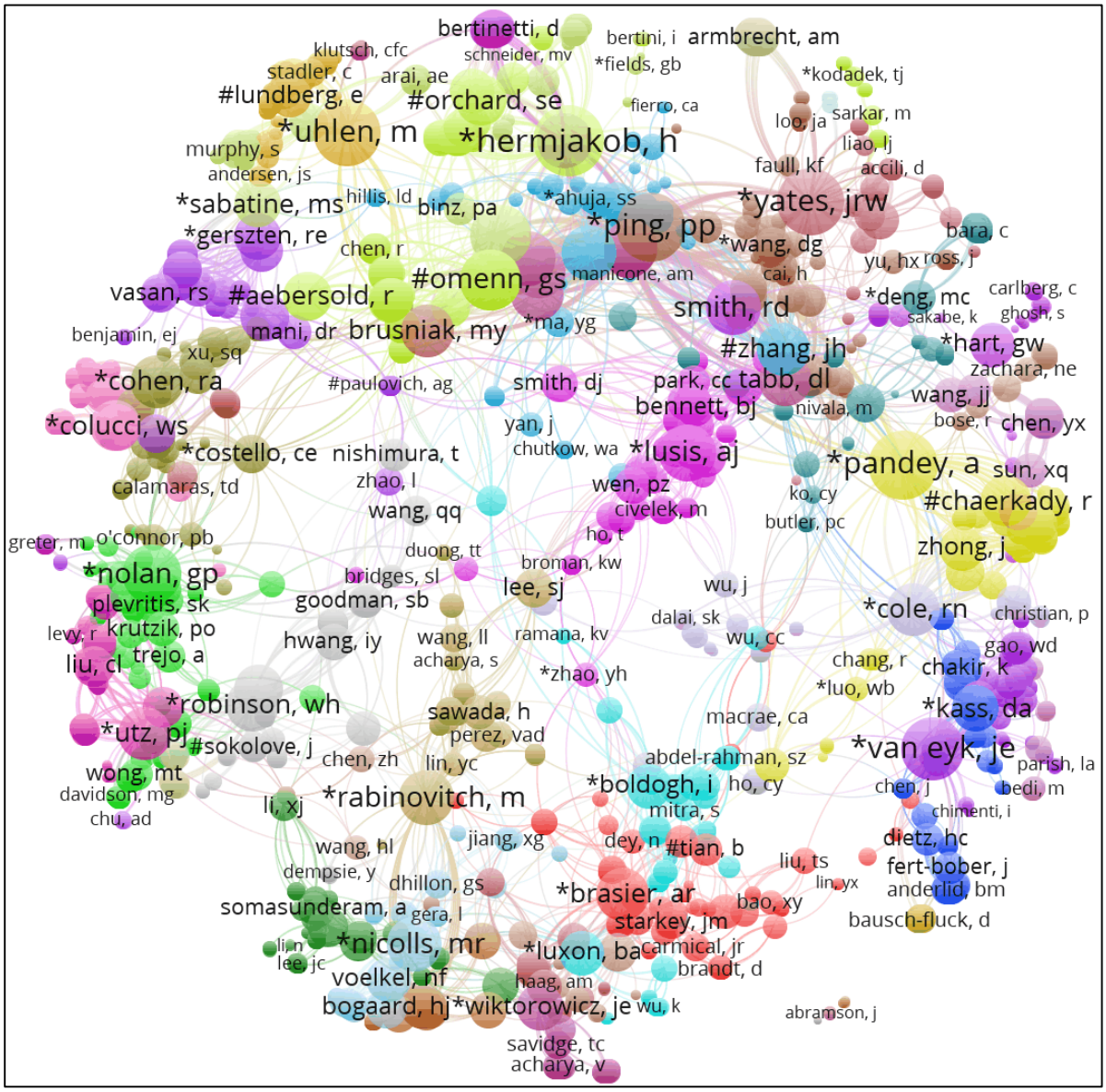


Figure 8. Co-authorship of Network NHLBI Proteomic Centers Publications: All authors 2010-2012.
Note: * = Center personnel; # = Outside Experts; ^ = NIH personnel

Publication Co-Authorship Network (2010-2013)

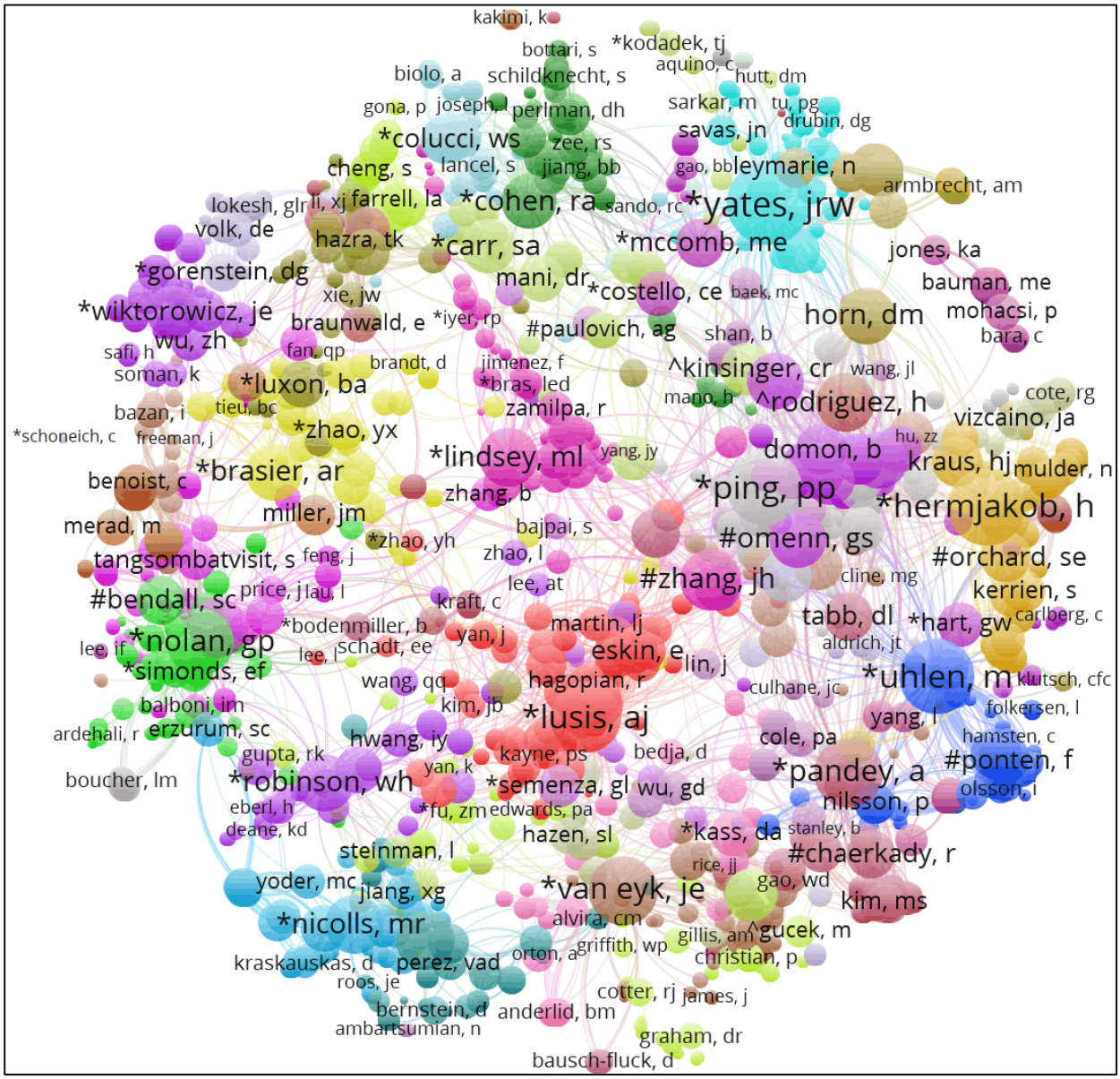


Figure 9. Co-authorship Network of NHLBI Proteomic Centers Publications: All authors 2010-2013.
Note: * = Center personnel; # = Outside Experts; ^ = NIH personnel

Publication Co-Authorship Network (2010-2014)

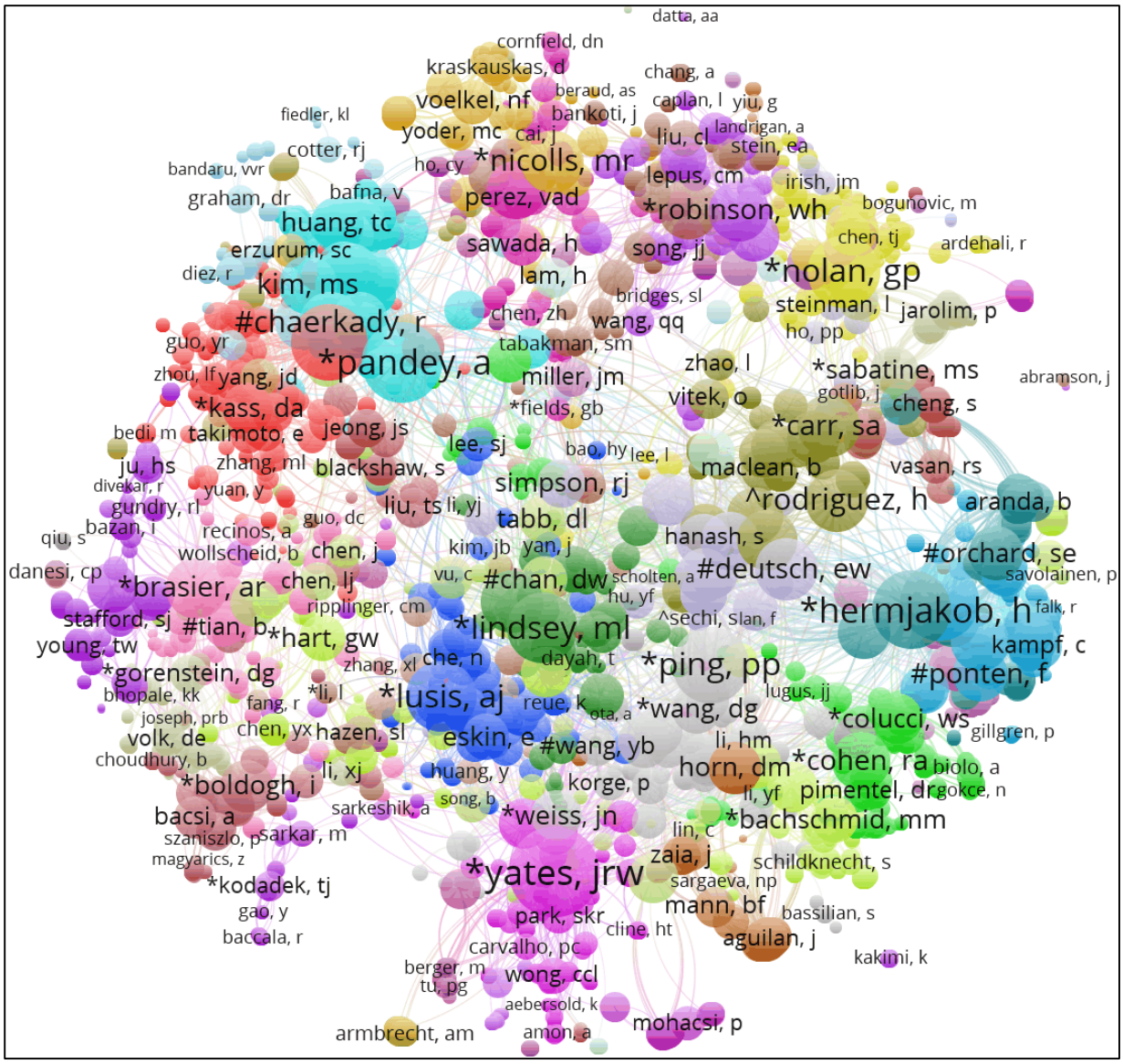


Figure 10. Co-authorship Network of NHLBI Proteomic Centers Publications: All authors 2010-2014.
Note: * = Center personnel; # = Outside Experts; ^ = NIH personnel

Publication Co-Authorship Network (2010-2015)

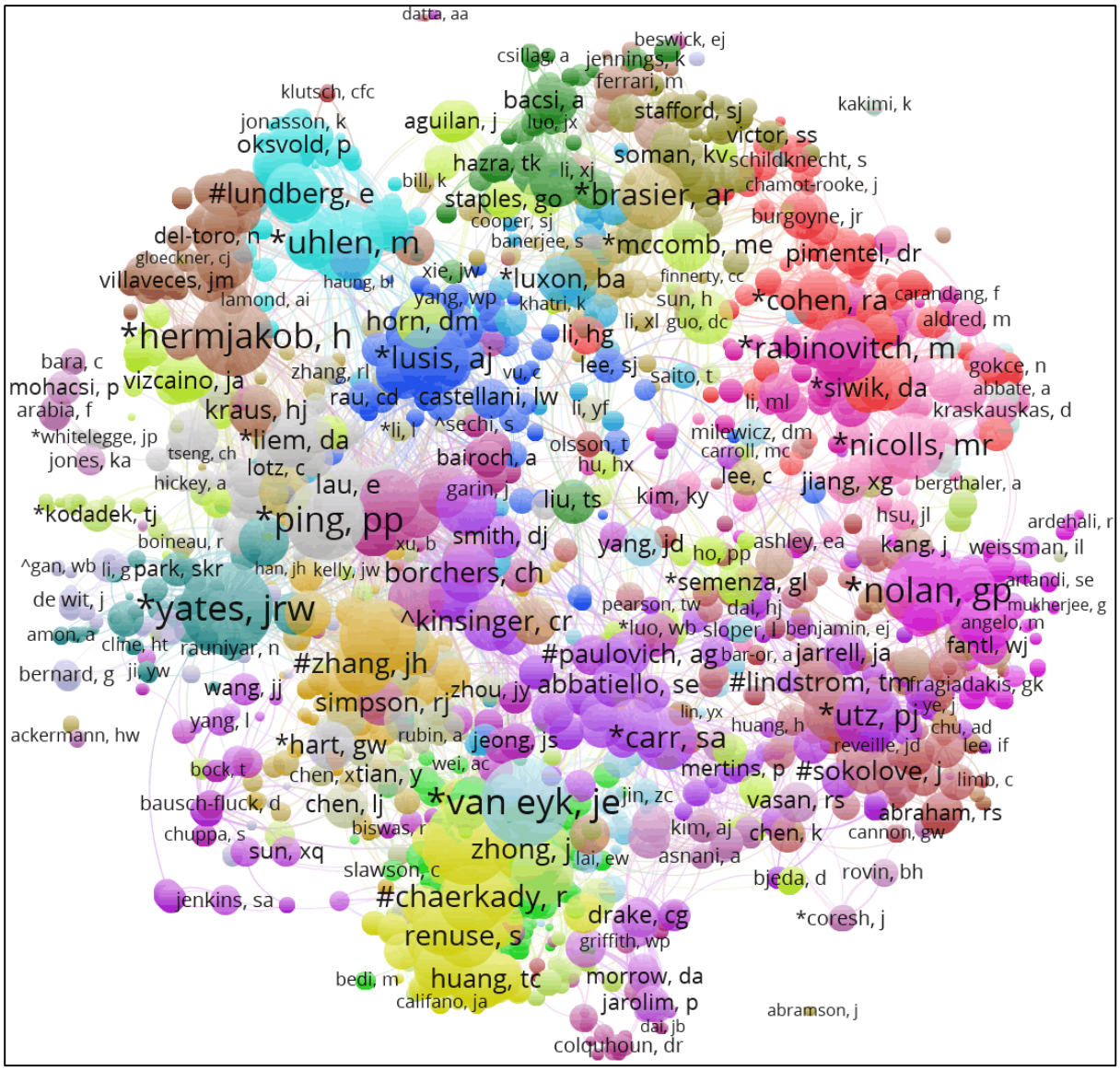


Figure 11. Co-authorship Network of NHLBI Proteomic Centers Publications: All authors 2010-2015.
Note: * = Center personnel; # = Outside Experts; ^ = NIH personnel

Visually, the expansion of the co-author network is clearly observed in **Figures 8 through 11 above**. Early in the funding cycle (**Figure 8 above**), the co-authorship network was fairly sparse, relatively speaking, with only a few prominently featured investigators and linkages present. However, as more publications and author teams are added each year, the co-authorship network expand dramatically with the network becoming more populated, dense, and thoroughly interconnected. This expansion is also confirmed in **Table 7 below**, as a substantial increase in the number of people in the measured network (i.e., nodes) as well as the number of links or ties in the measured network (i.e., edges), where a more that 8-fold increase in the authors and 12-fold increase in the connections are noted.

While there is an obvious structural growth in the network of co-authors functional relationships, as measured by the average number of ties between network members (i.e., average degree) and the average strength of the ties between network members (i.e., average weighted degree), showed a similar

increasing pattern. Indeed, the average number of connections between members, as well as the strength of those connections, increased each year as new members were added to the co-author network. Furthermore, the number of distinct communities of co-authors increased steadily during the period, despite the regularity in the closeness of the connections between members within and across these communities (i.e., modularity and density). Finally, the average number of steps it takes to get from one member of the network to another (i.e., average path length) indicates that by the end of the cycle, co-authors in the network were less than four times removed from one another. Ultimately, the NHLBI Proteomics Center personnel (the PIs, Co-PIs, Investigators, and Trainees) were part of a co-author network that was composed of more than 4,000 other researchers with more than 50,000 instances of shared authorship.

Table 7. Co-Authorship Network Metrics (by year)

Element	2010	2011	2012	2013	2014	2015
Nodes	530	1,329	2,183	3,205	3,864	4,379
Edges	4,009	11,204	19,050	33,132	43,008	50, 321
Average Degree	15.13	16.86	17.45	20.68	22.26	22.98
Average Weighted Degree	759.90	1,744.05	2,963.04	4,522.06	5,597.02	6,224.44
Density	.03	.01	.01	.01	.01	.01
Modularity	.94	.94	.94	.95	.94	.94
Communities	28	41	49	65	66	71
Average path length	3.04	4.89	4.36	4.05	3.80	3.69

From the interviews, respondents described many ways in which they collaborated with other researchers, both across Centers and outside of the Program. Cross-center collaborations, including informal information sharing across investigators, were described as increasing the quality of research and the pace of discovery. In terms of external collaborations, respondents spoke at length about the value of local, national, and international collaborations to their work, with some describing these collaborations as critical to tackling the scope and size of scientific challenges they faced. Respondents also valued their collaborations with industry, though, overall, this was discussed less frequently than collaborations with other researchers. Here, the primary themes were (1) the role of industry in taking proteomics technology to scale and t (2) the role of industry and a research partner in the development of new tools and technologies. Several comments from respondents illuminate their experiences with collaboration:

> *…the ability to transfer information from one disease - When you have Centers that work on the technology but work with clinicians in various different fields, the ability to transfer information is really exponential…And then the collaboration between groups that are focused on systemic manifestation of autoimmunity and cardiovascular manifestations, which is central to NHLBI, has been huge. –Investigator*

> *…on the clinical side, the technology, the techniques we used and developed led to collaborations with a couple of other groups we are making measurements with that were novel. –Investigator*

> *Somebody else presented data on epigenetics, and then [Name] and I realized that we could apply this using CyTOF, so we're generating all the reagents in this past year to apply it to CyTOF. –Investigator*

> *Even though they weren't formally a part of it, they created new pipelines for data analysis that we weren't necessarily aware of because we don't track the biostats literature.*

So that was the other thing that [Name] facilitated. If someone had new methods, [we] have them visit. –Investigator

Furthermore, the collaborative network of co-authors extended globally to include researchers in 50 countries. **Figure 12 below** indicated the geographic reach of the Program collaborations.

Geographic Reach of Collaborations

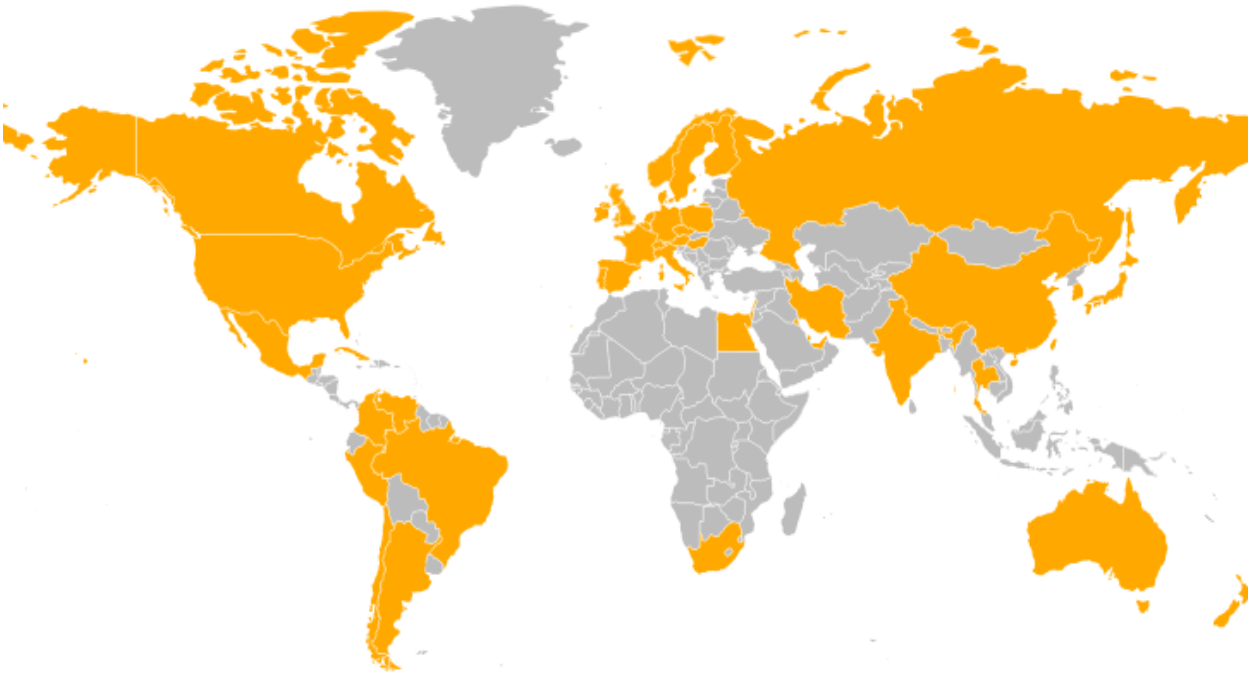


Figure 12. Co-Authorship Network Global Reach of Collaborations

According to those interviewed, the collaborations established as a function of the NHLBI Proteomic Centers Programs proved to be instrumental. Respondents from several Centers described how broad research networks had developed, many of them locally, around the work of their Centers. These networks grew, among other means, as scientists from other fields were attracted to integrating proteomics into their work. Respondents also discussed how the Program had contributed to their own personal networks and professional relationships, both by strengthening existing networks and fostering new ones. Within this, respondents saw value in how the Program was generating a cohort of well-connected trainees who understand the value of working collaboratively within networks. Some investigators described how the development of the science, in a broad sense, required extensive networks of scientists from multiple fields, and from multiple sectors. While some respondents described this as an existing need, others gave examples of how and where these networks currently exist and the role that the Program plays in them. This includes examples of how tools or knowledge developed through the Program were leveraged for use in other fields.

…like in our collaboration with [Person] who when I started thought proteomics was the stupidest thing ever. And now every single one of [Person]'s grants we are co-principal investigators on it and [Person] incorporates it in everything and understands the power. –Investigator

And by our last count, you know, there were some 30 or so R-series grants that included exclusive collaborations with members of the Proteomics Center or used proteomics technologies that wouldn't have done that five years ago. -Investigator

I would actually say [the collaborations were] quite critical. I mean on the one hand you definitely gain in a lot for your actual project - whether it is access to a technology that you don't have in your lab or samples that you can't easily acquire - but it is also just being able to talk to someone who's coming from a different perspective about your project and to really get your mind open to others that you hadn't considered before. –Trainee

…these collaborators are very well connected, and so there were cases I needed samples or different expertise, and the collaborators had good advice of who to contact to get that. Basically, that's having a good network. –Trainee

The Extent Science Informed by the Research was Conducted by the Program

We examined the extent to which subsequent science was informed by the research conducted by the Program. There was general agreement among respondents that, over the course of the Program, significant growth has occurred in the field. For example, the field has progressed in technical development, engineering research, and clinical advancement. However, there were differing opinions, particularly by individuals internal and external to the Program, about the extent to which the Program has contributed to these advances. Overall, investigators within the Program reported having made significant progress in the development of tools and generation of knowledge with clinical relevance, and they provided multiple examples of the clinical significance of their accomplishments. Some respondents also described the significance of the work to diseases beyond heart, lung, and blood. Centers within the Program also prioritized the dissemination of new research tools to the field, through various means and to various degrees. Publications were a common method of dissemination. Additionally, Centers disseminated tools through active outreach (including meetings and symposia) and by making tools and databases publically available online.

Respondents presented many examples of how tools developed through the Centers were being adopted and applied by other investigators and in other fields. While investigators from within the Program spoke of these tools as being, in some cases, highly influential, assessments of respondents external to the Centers varied. Respondents described how the opportunity to work closely with other investigators from other fields expedited the process of discovery in the Program, as it facilitated communication and exchanges that accelerated the advancement of knowledge and understanding.

Usually people are interested. A lot of times they're people that are not in the proteomics themselves. They're in maybe stem cells of something like that and they want to be able to monitor protein expression or something like that in different stem cells or with different differentiations. –Trainee

Well, the pipeline that we developed and in large part developed through this heart, lung, and blood program is basically the pipeline that the proteomics community is now using for analysis. So it's been very influential. –Investigator

It [the Program] really opened up a new area of research for the entire lab…it's a really interesting story that we wouldn't have necessarily pursued if we didn't have the protocols and everything that we had set up for the proteomics grant. –Trainee

I think this program has driven the scientific community since the overall inception of the first part of the Program. There was a lot of interest within the scientific community to study cardiovascular disease at the molecular level. And then more specifically, we saw a huge uptake in research being focused in the area of applying proteomics technologies as they matured to the study of cardiovascular disease models. –Investigator

Furthermore, to assess science informed research in the field, an important part of advanced bibliometric performance and impact evaluation is the construction and representation of a research profile for a specific entity.[14,15] A research profile is a breakdown of output, performance, and impact according to internationally defined research fields on the basis of the journals used by the entity. In analyzing the profile of interdisciplinary research, and computing the normalized metrics for worldwide comparison, we focused only on those papers considered "research articles".

The largest mean citation score (i.e., per paper average) values were found in the fields of *Multidisciplinary Sciences* (MCS = 131.25); *Medicine, Research and Experimental* (MCS = 41.75); and *Cell Biology* (MCS = 33.85). This suggests that the NHLBI Proteomic Centers research is generating substantial interest in these disciplines. However, in order to fully understand the relative differences within and across fields, an examination of normalized (i.e., adjusted for age, type, and source) metrics of worldwide impact is necessary. **Figure 13 below** provides a spectral analysis using normalized indicators of the research output from the NHLBI Proteomic Centers across those fields with **3 or more publications**. The color of the shaded bar corresponds to the average field normalized score for the set of papers published within the respective field.

The field normalized citation scores (MNCS) indicated that publications in these fields were highly influential and visible, far exceeding the number of citations expected for publications of the same age and type in the respective fields. Even in those fields where relatively fewer papers were published, the measured impact of those papers was still high. Results revealed that of the 47 fields with **3 or more publications**, the largest normalized impact was seen in the fields of *Multidisciplinary Sciences* (MNCS = 7.46); *Medical Laboratory Technology* (MNCS = 6.63); and *Medicine, Research and Experimental* (MNCS = 4.65).

Although these fields had large normalized values as a measure of impact, all three had relatively few papers associated with the field, signaling the presence of a small number of very influential papers amplifying the value. In contrast, the largest volume of papers was associated with the fields of *Biochemistry and Molecular Biology* (n = 207; MNCS = 3.01); *Biochemical Research Methods* (n = 192; MNCS = 2.47); *Cell Biology* (n = 111; MNCS = 3.37); and *Cardiac and Cardiovascular Systems* (n = 108; MNCS = 2.31). The MNCS for these four fields were still far above average (i.e., > 1.5), indicating the papers in these fields far exceeded the number of citations expected for publications of the same age and type.

Furthermore, the total number of citations accounted for by these four fields was 13,830, with 5,827 citations in *Biochemistry and Molecular Biology*; 2,395 citations in *Biochemical Research Methods*; 3,757 citations in *Cell Biology*; and 1,851 citations in and *Cardiac and Cardiovascular Systems*. Collectively, the recognition of work in the four fields by other researchers worldwide reflect the focus of the NHLBI Proteomic Centers in applying proteomic approaches to gain a better mechanistic understanding of the physiologic pathways underlying defined clinical conditions related to heart, lung, and blood diseases.

[14] Van Raan AFJ (2008). R&D evaluation at the beginning of a new century. *Research Evaluation.* 8: 81–6.
[15] Van Raan AFJ (2005). Measurement of central aspect of scientific research: performance, interdisciplinary, structure. *Measurement.* 3: 1–19.

Research Output from the NHLBI Proteomic Centers Across Fields

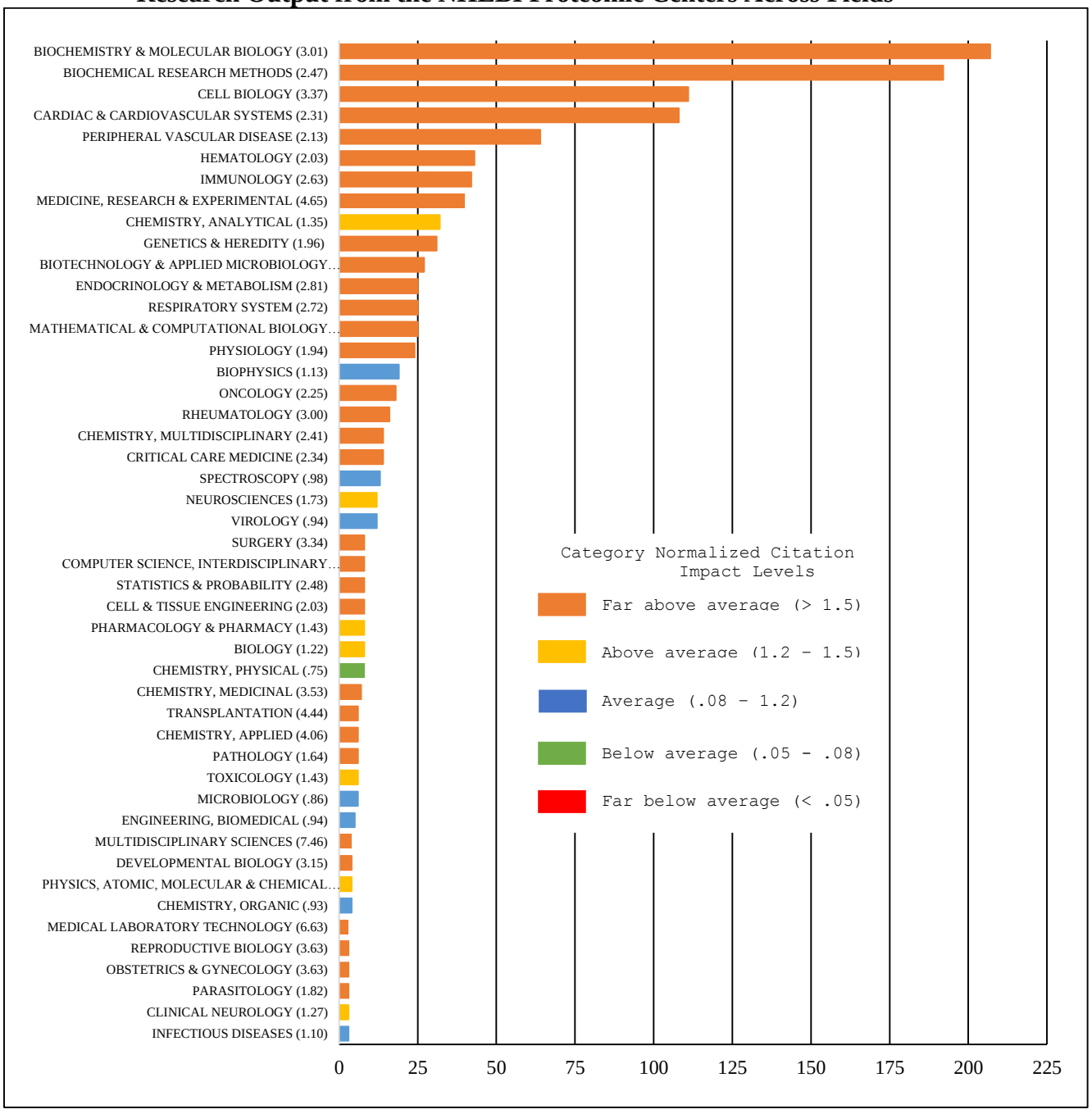


Figure 13. Research Output Profile for NHLBI Proteomic Centers with MNCS in Parentheses

A secondary analysis of the top 1% highly cited papers from the NHLBI Proteomic Centers Program illuminated the scientific fields that were citing this highly visible set of research output. The 52 papers accumulating enough citations to be in the top 1% of all scientific articles found in the WoS database had been cited 7,952 times by other researchers. This accounts for 38.9% of the total number of citation. Thus, this set of 52 papers was a highly recognized group of scientific output produced by the Centers, from which we were able to examine the range of fields the set was subsequently influencing. In **Figure 14 below,** the 52 papers were published primarily in the fields of *Biochemistry and Molecular Biology*,

Biochemical Research Methods, *Cell Biology*, and *Biotechnology and Applied Microbiology*. Those researchers citing these 52 articles from the Centers were publishing in those fields as well, but also in the areas of Oncology, Genetics and Heredity, Pharmacology and Pharmacy, Neurosciences, Hematology, and Analytical Chemistry. In all the 52 articles, the 1% level papers were published in 10 fields. Subsequently, these 52 papers were cited by other researchers in papers published in 109 different fields. Thus, the NHLBI Proteomic Centers Program was subsequently influencing a range of scientific disciplines and fields worldwide.

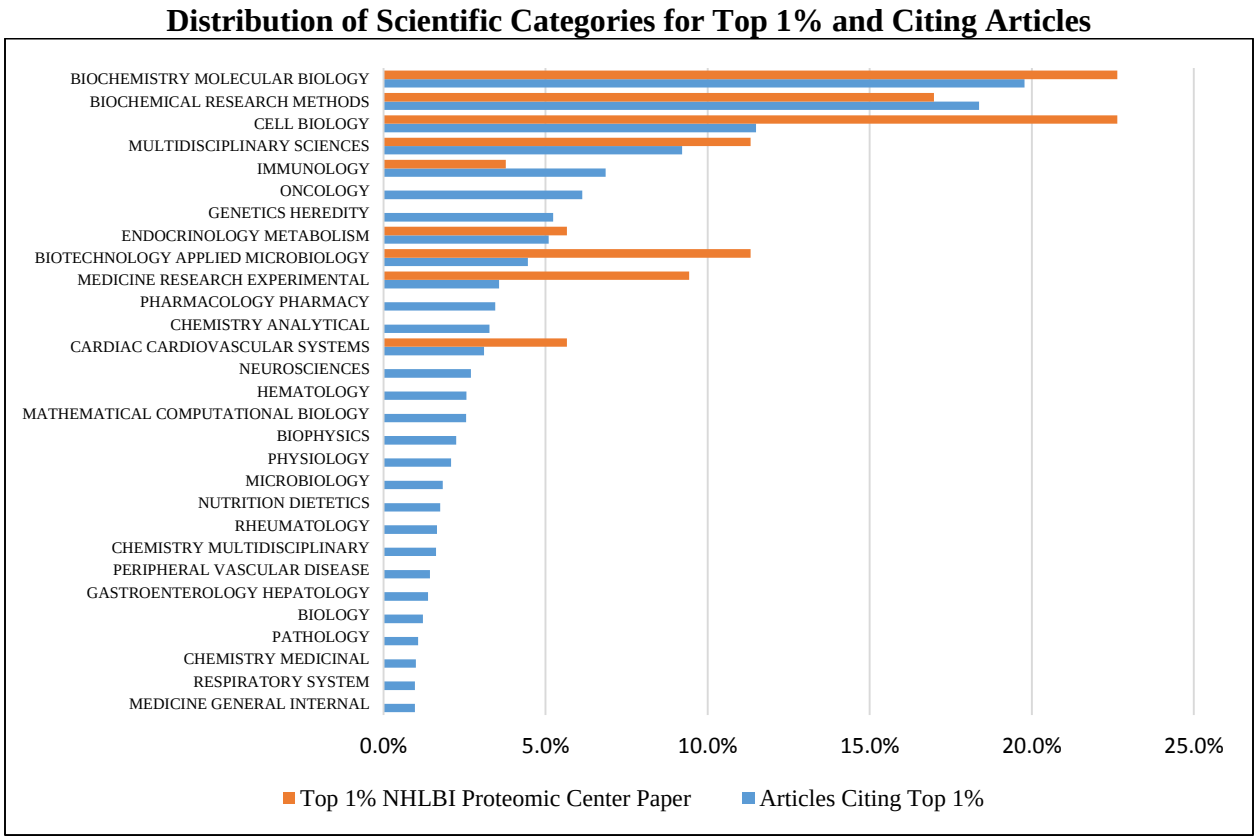


Figure 14. Distribution of Scientific Categories for Top 1% and Citing Articles

The Extent the Program has been a Springboard for Subsequent Academic Appointments and Professional Recognition

Overall, the Program fostered an environment within which many stakeholders felt supported in the development of their careers. In particular, both current and former trainees spoke very highly of their experience with the Program with regard to their training, mentoring, and advancement. Stakeholders from multiple groups described how trainees benefitted from opportunities to work and gain expertise in multiple disciplines. This was facilitated within the Centers through opportunities to work as part of interdisciplinary teams and to work in the labs of multiple senior scientists. At the program-level, stakeholders felt supported through cross-center exchanges and training opportunities. In addition to gaining specific skills through the Program, trainees described how they had been intentionally and skillfully mentored by senior scientists, and how they had benefited from a range of professional development opportunities including presenting their research at major conferences, managing labs, and overseeing junior trainees.

Respondents also commented on how opportunities to collaborate outside of their Centers had contributed to their training. As for many trainees, their experience with the Program translated into opportunities for professional advancement, and several former trainees described ways in which they were shaping the field in their new positions. Trainees also described how their opportunities had expedited their training. For example, they commented that they were able to move into other positions more quickly than other trainees in the field.

They also noted several hindrances as well. Hindrances to the training of junior staff were also described, primarily in terms of isolation experienced by some trainees. This was described at the center-level (e.g., trainees isolated to their own project, and not well integrated into the Center) and at the program-level (trainees isolated within their Centers, with limited opportunities to meet or work with members of other Centers).

The extent to which the Program served as springboard for academic appointments was examined primarily through reported information on the biosketches and secondarily though information derived in the interviews. Several instances of position advancement were reported by those considered key personnel in the Program. Overall, 42 separate advancements occurred during the period under inquiry. This was reported by 33 or 44% of key personnel who submitted biosketches or CVs. **Table 8 below** shows the distribution of position advancements by year across three career "tracks" (academic, research, and medical) identified in our analysis.

Table 8. Position Advancement Across Career (by year)

Level	2010	2011	2012	2013	2014
Academic Track					
Assistant Professor	1		3	1	
Associate Professor	2	1		2	4
Faculty Member		1			
Professor	2	4	2	2	1
Named Professor		1		1	
Research Track					
Researcher				1	
Scientist		1			
Associate Director			1		
Director or Co-Director	1		1		1
Medical Track					
Fellow	1		1		
Assistant/Associate Physician					
Physician (Attending)				1	1
Chair/Chief	2		1		1
Total	**9**	**8**	**9**	**8**	**8**

We found that advancements occurred across the three tracks for every year during the period of inquiry. We also determined that advancements occurred across levels within tracks, indicating that advancements were taking place even at the higher levels of academia. It should be noted that 2015 was not included in our analysis since the biosketches/CVs were submitted to the evaluation team early in the final year of the Program. Nevertheless, the consistency of advancements every year suggests it would be reasonable to expect some advancement would have occurred in 2015 as well. In addition, several of those interviewed,

highlighted instances where they recognized position advancements in the context of the Program and its activities, further supporting the pattern reported above:

> *The three of us who are here started in 2010 or 2011 and we just got it…all three of us now have faculty positions. Actually we have three that are assistant professors now. –Trainee*

> *In the interval period since this contract started, I've moved up to be the Director of our M.D. Ph.D. program here, and I think the fact that I've been able to get some of these large program project grants and to manage them has been really helpful. –Investigator*

Furthermore, others referenced the effect their engagement in the Program had on their career development:

> *I was assistant professor when I wrote the application…and a week before it was due, I found out my promotion had gone through. So I was becoming an associate with tenure and I put that in the application. I'm an assistant now, but effective [specific date] I will move to an associate…it's been a huge increase in my career. –Investigator*

> *If I had stayed in basic research, there's no way that in four years I would have become an assistant professor. It advanced my career very quickly…I was able to publish a lot, get collaborations, get to know people, get funded, get promotions… -Trainee*

Furthermore, in terms of professional recognition, individuals affiliated with the Program were recognized both nationally and internationally. For example, individuals were published in high-impact journals and invited to speak at both national and international conferences and meetings. In terms of understanding the extent to which the Program served as a springboard for professional recognition, we examined several aspects related to the concept of recognition. First, we noted specific honors and awards in several categories reported by key personnel, as attention to the achievements and status of those in the Program. Second, we looked at the scientific recognition, as evidenced by the citation counts and normalized impact indicators of scientific publications. Citations symbolize the association and acknowledgement of scientific ideas, and the references which authors cite in their papers make an explicit link between their current research and prior work in the scientific literature archive.[16] Thus, we determined that citations, as the exchange and recognition of research output, is desired by researchers and one of the key driving forces in the advancement of scientific work. Finally, we examined the acquisition of funding by key personnel as be a measure of success. In this case we assumed that subsequent funding was in some way the recognition of scientific quality and capability promulgated by those associated with the Program.

Specific individuals in the Program were identified as having influence and stature both within the Program and in the broader field of proteomics. Acknowledgement and recognition on the basis of honors and awards was observed within multiple categories, as shown in **Table 9 below**. Furthermore, we examined the instances of recognition within categories by the years of experience of the key personnel included in the biosketch/CV sample. We found that across categories, instances of recognition were not limited to those with advanced careers. Nevertheless, those with advanced careers (i.e., more than 20 years) noted a substantial number of NIH funding awards (including subprojects) over the five-year period.

[16] Koskinen J, Isohanni M, Paajala H, et al. (2008). How to use bibliometric methods in evaluation of scientific research? An example from Finnish schizophrenia research. *Nordic Journal of Psychiatry.* 62:136–43.

Table 9. Acknowledgements and Recognition of Key Personnel

Categories	0-10 years	11-20 years	> 20 years
Fellowships (n = 63)	2	15	15
Awards (n = 63)	7	26	37
Distinguished Professorships (n = 63)	0	2	8
Lectureships (n = 63)	0	5	9
NIH Funding Awards, by year (n = 59)	27	243	798
Non-NIH Funding Awards, by year (n = 48)	30	77	77

Several interviewees described their experience in the Program relative to the professional recognition accumulated during the Program period:

> *The Program itself has helped my career and my recognition in the national and even in the international capacity. –Investigator*

> *I have really benefited from the Center. When I first started in my other labs, my curriculum vitae was almost empty…I only [had] a couple of publications… I mean almost nothing. And, I didn't have first author publications when I joined this lab. But now, I have more than 35 publications, recent publications and more than 15 abstracts and 3 book chapters – one is first author. And I've got a couple of awards. –Investigator*

> *We were provided the opportunity to go and present our work. So, while I belonged to the Center, I believe I went to a minimum of two meetings per year, and I always presented in the meetings - both doing oral presentation and posters…and my colleagues were doing the same. -Trainee*

> *We all give lectures in our own institution and other places as well. Those often end up with people coming to talk with you afterwards [about] research projects as well as responses to papers that we publish. –Investigator*

In our examination of the scientific recognition on the basis of citations patterns, we found the cumulative scientific output of the Centers to be substantively recognized by other scientists. **Table 10 below** lists the performance and impact metrics for the set of NHLBI Proteomics Centers' publications by total and by year.

Table 10. Performance and Impact Metrics (total and by year)

	Papers (P)	Total Citation Score (TCS)	Mean Citation Score (MCS)	Mean Normalized Citation Score (MNCS)	Mean Normalized Journal Score (MNJS)	MNCS/MNJS	I3
Total	975	20,414	20.93	2.72	1.42	1.92	
2010	71	3,568	50.25	2.89	1.55	1.86	225
2011	162	6,336	39.11	2.68	1.20	2.23	472
2012	200	4,265	21.33	2.71	1.54	1.76	562
2013	229	4,249	18.55	2.77	1.41	1.96	683
2014	191	1,874	9.81	3.20	1.42	2.25	512
2015	122	122	1.00	1.85	1.52	1.22	253

Our analysis included normalization of citation data, controlling for journal, field, and year of publication in order to facilitate comparisons to all indexed scientific literature. As a benchmark, a ratio of 1.0

indicates that the set of publications have met the expected number of citations based on the journals in which they were published. For the collection of 975 papers, the MNJS ratio was 1.42, indicating the papers from the NHLBI Proteomics Centers were cited 42% above what was expected across the set of journals publishing papers produced by the Centers. The MNCS ratio was 2.72, indicating the papers from the NHLBI Proteomics Centers were cited 172% above what was expected worldwide across relevant fields.

In terms of evaluating the level of impact, generally accepted international impact standards for interpreting both the MNJS and MNCS ratios have been published. Specifically, the levels are: far below average (indicator value < 0.5); below average (indicator value 0.5 – 0.8); average (indicator value 0.8 – 1.2); above average (indicator value 1.2 – 1.5); and far above average (>1.5). Thus, the impact of the network papers within the published journal set was above average and the impact of the Centers' papers in the fields was far above average.

We further evaluated the global standing of the journal set containing the papers from the NHLBI Proteomics Centers relative to the fields in which the journal belongs. The MNCS/MNJS ratio was 1.92, indicating the mean citation score of the network's journal set exceeded (by 92%) the mean citation score for all articles published in the fields to which the journals belong. Thus, Center researchers as a group publish in journals with a high impact in the fields of study with relevance to proteomics. **Table 10 above** also includes performance and impact metrics by year, and indicates 2014, has the highest performing and impactful set of papers based on the metrics calculated. Based on the integrated impact indicator (I3), 2013, had the highest overall impact based on the weight of highly cited papers published. Thus, 2013-2014, appears to be the highest performing and impactful two-year period for the NHLBI Proteomics Centers Program.

More generally, the work being conducted by individuals was viewed as important and necessary in the field of proteomics, especially when discussing biomarkers for cardiovascular disease. This was evident in additional funding leveraged by these investigators, and suggested a need for future funding for the expansion of the knowledge being created. The linkage of key personnel to subsequent funding during the period under inquiry was noted through our examination of NIH funding as archived in NIH Research Portfolio Online Reporting Tools (RePORT). We compiled the funding for 57 key personnel who had a funding record in NIH RePORT. We identified the awards made to the 57 key personnel prior to and during the period of inquiry. As shown in **Figure 15 below**, the computed linear trend for the period ***prior*** to the evaluation timeframe of 2010-2015 was lower (2005-2009). The difference in the linear trends suggests the ***level*** of funding during the period of inquiry increased during the Program from identified funding prior to the Program.

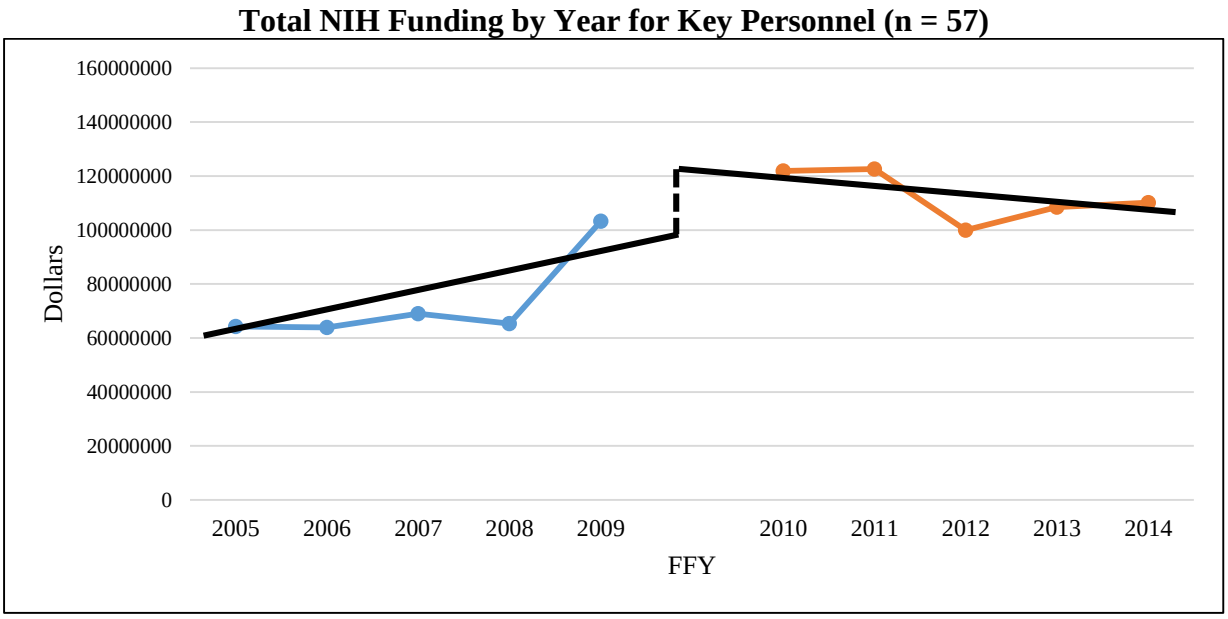


Figure 15. NIH Funding for Key Personnel (by year)

Similarly, we examined the average funding award level by NIH institution prior to and during the period under inquiry. As shown in **Figure 16 below**, the average award size increased in several institutions, including National Institute of General Medical Sciences (NIGMS), National Center for Research Resources (NCRR), National Cancer Institute (NCI), National Center for Advancing Translational Sciences (NCATS) and National Institute on Alcohol Abuse and Alcoholism (NIAAA). Other institutions – National Institute of Biomedical Imaging and Bioengineering (NIBIB) and National Human Genome Research Institute (NHGRI) – saw a reduction in the average award size during the period. However, this would be expected as the work of the researchers who may have previously received funding for genomics research likely shifted their focus to proteomics during the period.

Furthermore, the biotechnology research engaged in by Center researchers, previously funded by NIBIB, was likely supported by NHLBI Proteomic Centers awards during the period. Nevertheless, the total average award size across all institutes and Centers for all Program investigators increased during the period despite the shifts among the different institutes and Centers.

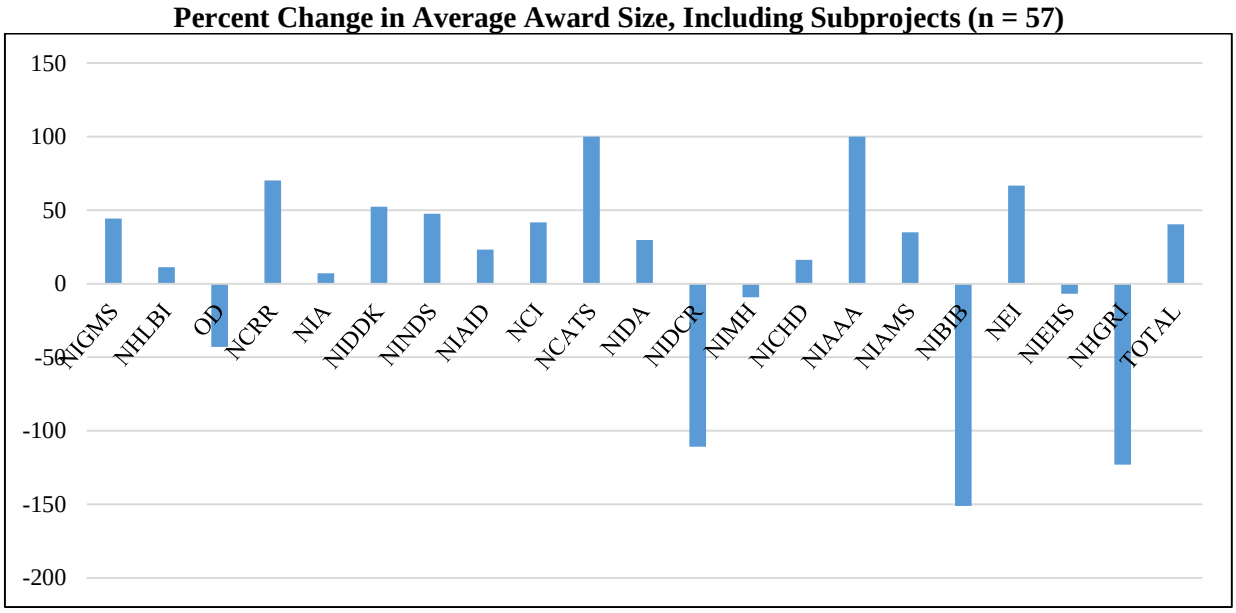


Figure 16. Average Award Size (by percentage)

From the interviews, respondents indicated their experience in the Program was connected to acquisition of additional funding for their work:

> *Also, from a career standpoint, I've been able to get other big program project grants with [Person] and [new Person] and other people… -Investigator*

> *For me in particular this work has been so critical in advancing my own career. I was on a training grant, [a] training grant that was about to expire later this year and I applied for several different NIH and foundation grants using the Proteomic project as the central thesis of all of these grants…I was lucky enough to get two of these grants very recently. –Trainee*

> *A lot of the things that I discovered in the proteomics field is…I'm writing on grants now. So I've already pitched an idea to [Person] and we are going to try to write a grant together. –Trainee*

> *One of the grants that I got was actually, you know, it was told to me kind of discretely that my grant actually scored the highest in North America. It was very, very well-received. –Trainee*

While it is not clear as to what, if any, influence the Program and activities had in terms of directly leading to the success of those involved in the Program, it is clear that stakeholders recognize their work related to the NHLBI Proteomics Center agenda was influential to how they were recognized and acknowledged as professionals.

VI. Outside Expert Perspective

In our evaluation of the extent to which the Program met its objectives, we determined that the Outside Experts constituted a valuable source of information. This enabled us to contrast what we heard from those within the Program. Thus, in our final assessment, we treated the key informant interviews from the Outside Experts as a separate data source in addressing the evaluation questions. Below are the results of the analysis of their interview data and responses provided as examples to illustrate their perspective relative to the evaluation questions. All of the direct quotations in this section are from those interviewed as Outside Experts.

The Extent the Program Developed and Shared New Tools and Technologies

In general, Outside Experts voiced positive assessments of the new tools and technologies developed through the Program. This finding is similar to the data represented from the rest of the stakeholder groups, but comments from the Outside Experts were general, while other cited specific tools, technologies, or methods that emerged from the Centers.

> *They had a really nice paper maybe three or four years ago that had major methods and major protocols…it looked at cardiovascular problems and issues. I was like, "Oh, wow, I didn't know you could do that"…*
>
> *I think in terms - particularly in the application of mass spectrometry - there are a number of software tools, as well as collaborative developments with instrument manufacturers, that are having an impact.*
>
> *Absolutely…no question. Now, I'm not a clinician, I'm a basic scientist. I look at the data that [Person] is generating and look at how that can signal individual cells. And that's very important. [Person] changed our way of thinking.*

One respondent, in contrast, offered a differing assessment of the new tools developed, while another respondent highlighted a continued need for additional tools and technologies.

> *On the other hand, with the exception of the [Center] group, how many of these people are really true tool developers? So, you might think if they were going to renew this [grant], they might think about something more in tool development groups. Because, generally, those groups also know how to get the data out there and the tools out there…I guess I would love to see a little more tool development going forward.*
>
> *We're missing tools in our toolkit. The average end user laboratory doesn't have the time, money or expertise to go develop these assays. A community resource needs to be developed and supported to enable everybody democratized proteomics. I think that's the biggest hole in moving this stuff into the clinic. First thing they are going to do is go to the commercial catalog and find the few assays that everybody else is also using – that is what we all studied because that frozen assay exists. We are missing a critical resource, a critical tool in our tool kit, which is off-the-shelf assays for clinically-relevant human proteins. I think before this field can move forward substantially that is going to have to get built.*

Among some Outside Experts, there was a recognition of tools and technologies developed by investigators involved in the Program, but a lack of clarity around whether or to what extent the Program supported this work. Furthermore, a lingering question was also to what extent the development of these tools was attributable to the Program.

There are some technologies that have come up frequently, the technology that's come out of [Person]'s lab was not an outcome of this particular consortia, right? It involved [Person] who invented [technology], that's a major advance, but as far as I know, it was not developed specifically as part of this Program. It was part of the [Person]'s research program. [Person] had been working on this for many years.

So, yes, we have been very aware of what has been happening and using the tools that [Person] is developing…indeed, I am aware of all the tools that are being built, but I don't know who is funding them.

Several Outside Experts described how the Centers shared their tools and technologies. The most cited method was through web-based tools, though some respondents also described how dissemination had occurred through workshops, personal relationships, and connections with Center investigators.

Interviewee: I've seen them [tools] and I've played with them a bit from some of the websites and things, but my work isn't far enough along to use those tools.
Interviewer: You mentioned websites, are there any other channels that you were able to access new knowledge and tools developed through the Program?
Interviewee: I've only looked at the web-based tools…Very impressive.

Mostly from the website…this is the beauty of the [Center]. I think they actually built a very strong interface with the community with their website. There are many, many online tools that you can download, thousands - I have heard about tens of thousands of downloads already - that's a far-reaching effort.

The other side is the [Center] also became a source of information, technology development, and expertise, where they provide seminars. They provide workshops that become very valuable to the local community.

One respondent, in contrast, implied that the Program could have done a better job in terms of sharing tools and technologies.

So it might have been easier to know more about the Program. Just now, I went to their website and I'm reading about it - it's like "our latest news", but other than looking in journals, that's the only way I would find out about what's going on. Now that I'm on their website, they do have a newsletter, so I probably should've signed up for their newsletter or maybe they should be more active in sending out their newsletter to individuals…that might be beneficial.

Several respondents suggested that even when useful tools are developed and shared, they may be cost-prohibitive or otherwise inaccessible to much of the field.

I think the cost of a lot of these technologies is in excess of what a lot of people in smaller labs can afford. The level of complexity it takes a small army to run a lot of these technologies are not accessible unless you are at a place where you can work with someone to do that.

…the way I see access is not just the tool itself, but also improving those tools for a diverse background of expertise. That is going to be the challenge. So you have to kind of dumb down the interface, almost like turning a very sophisticated computer into a Mac. How many people understand a Mac or an iPhone? You don't need to know necessarily…all you need know is how to push buttons. I hope they're moving towards that direction.

…just looking at the caps of how it's funded, these are like the top labs. The top analytical capabilities and also the most resources. I look at it from the government perspective, I would say, if I was a smaller lab these labs can do it because they have tons of funds and tons of expertise, but how would I apply that to my own research? So it's not necessarily the methods out there and the publications out there…but at the end of the day, the implementation is quite expensive. If you don't have grants, university support, or funding to buy the equipment, it's a different story.

The Extent the Program Created Outputs of Clinical Utility

None of the Outside Experts directly stated that the Program had produced outputs that had matured to the point of clinical utility, and some commented that this was not a feasible goal given the Program's scope. However, a common perspective on this point was that the Program was "*laying the groundwork*" for clinical applications, but that they were not there yet. One respondent suggested that the work of the Program was just beginning to impact clinical work.

I think things like proteomics and genomics are typically more powerful tools for discovering and understanding disease. They are not the kind of thing you would necessarily use directly in patients. Although there is some in use now – for example, genomic sequencing for cancer – typically speaking, they are more tools to understand the biology. Having said that, though, I think some of these groups, like [Center] group, for example, and [Center] group, are trying to identify signatures that correlate proteomics signatures to specific diseases. And what I don't know is how far they've taken that to actual clinical applications, but certainly the implication there would be these signatures could eventually be used for clinical application.

…[Person] has made significant improvements in the understanding of the field and has provided tools to help other people. That is the only program I'm really intimately aware of. I can't really comment anymore other than that.

Interviewer: Do you feel that that is leading to clinical outcomes yet?
Interviewee: No, not yet, but I think it will in the future. [Person] is laying the groundwork and the tools and developing what will certainly have a long-term future, I think. Maybe a knowledge resource.

By setting up methodologies, where you can easily determine potential targets of interest, you're saving tons of time and money - that was a finding coming out of the proteomics core Centers. They basically said, by using this multiplexed approach, we can say that these are the things that should be followed and that we should spend more time doing. So as opposed to just using a shotgun approach, we can use a semi-targeted approach and find out what things are of most interest, and then from there, decide what we should build. It helps dwindle down the candidate list. So I think that advance in the methodologies and what they showed is really powerful for the entire field. That is really getting close to how do you translate findings that are in a lab and then apply it in a more clinical setting.

I think the development of knowledge or tools relevant to clinical or translational questions been very important… So, you go in an emergency room — hopefully not – with a heart attack. They try and minimize the damage done by that. In the past, they have used primarily biochemical methods or drugs, pharmaceutical ones. Now, because of studies like [Person] is doing, they are thinking of different approaches and using different approaches.

Other Outside Experts had a less positive perspective, indicating that the clinical utility of program outputs was limited.

> *I'm not really sure what else is going on in the translational research area for proteomics. It's kind of hard to say. Just a few examples that were presented at meetings.*

> *Again, it's [the Program] resulting in a lot of descriptive studies, which may or may not have bearing on the clinical disease…and that really has no bearing on the mechanism underlying the clinical disease.*

> *Interviewee: After all this discovery work, there aren't any new biomarkers that are approved by the FDA for clinical use yet. So you can't actually declare success at this point.*
> *Interviewer: Because it still needs to go through that round of approval?*
> *Interviewee: That's right, and the probability of success is very low. There are literally tens of thousands of publications of discovery of biomarkers. And out of those tens of thousands of publications, none have been clinically validated… So there's obviously a problem.*

Outside Experts described the push toward the integration of proteomics into clinical trials as something that still needs to happen, and several respondents implied that proteomics is lagging behind other areas (i.e. other -omics and cancer research) in this area.

> *I think we're getting the tools to the point where it's kind of moved things more into the clinic and not just more "methods of development". The methods we have, now what can we show and what can be done? I would imagine the way forward would be to put more of the chips into all right places. These are clinical trials, let's add proteomics on to the base of things we've learned in the previous five years...*

> *I think that if it compares with other -omics that are out there, especially genomics, I think genomics is still ahead just because it's better established. They have come up with actionable information, driving some of those discoveries into drugs… but proteomics still lags behind. It has to still has to prove itself. It has to materialize some of the discovery work they have been doing the last 15, 20 years…try to implicate those discoveries directly into actionable items - like by markers are drugs…somehow use the data in a more direct fashion. So, in that regard, I think technology-wise it's improving in getting much better. But to my knowledge, there hasn't been any significant discovery, a large high-impact discovery that directly comes from proteomics.*

> *I come back to the sort of paradigm that tools like proteomics are good for generating hypotheses. Then what you want to do in the next phase is test those hypotheses, validate them, and then use them…I don't know what's happening at in the heart, lung, and blood area as well - I know the cancer area a little bit better. So I don't know how far along these guys have taken their discoveries, but I suspect that is the direction they all want to go.*

In general, two respondents suggested that proteomics research to date has been too broad, and that progress toward clinical utility required the field to focus and integrate their efforts.

> *…from what I've seen, what would benefit the clinic, would be a much more focused approach on the biological pathways…a few pathways rather than looking at the entire realm of proteomics. I think we've collected a lot of information broadly and need to go more in-depth and combine across, not just proteins, but gene expression, metabolites, et cetera. Integrate these different types of data.*

What is necessary is, on the technical side, is to pay more attention to specific directed assays and to do a really precise job of measuring the most promising candidate biomarkers and then to apply those assays – by a large clinical staff or cohort. Working on directed assays and specific biomarkers hasn't been as popular in proteomics as doing just their broad discovery because it's a different technology approach and it's not technologically created, but it's necessary.

Commentary from Outside Experts on the contributions of the Program in the form of inventions, patent applications, spin-off companies, and licensing agreements was fairly limited. However, Outside Experts did comment on the value of these collaborations with industry in a general sense, with one respondent suggesting that this is a "*weak link*".

The people who have done well have actually worked closely with companies and have had a lot of hardware. That's where the best hardware people are these days, they really are in those spots.

I think in terms of - and particularly the application of mass spectrometry - there are a number of tools, software tools, as well as collaborative developments with instrument manufacturers that are having an impact there.

I think what's critical is getting something through the system so that people can see it work. To the best of my knowledge, I don't think anything has got clinical approval that's been identified by proteomics. I don't think there's been any biomarkers out there that have made it through and received full regulatory authority, approved in any country - FDA or anywhere. So I think achieving that would be the thing that would then give the impetus for everybody to get really focused on what they're doing. Because I think it was a bit of a hiatus at the moment. Everything looks great in the laboratory, but either they've been identified by small academic groups and have the funding to push it through further or the drug companies…and I think that link is still quite weak.

The Extent Research Contributed to Knowledgebase Linkages

In this area, again, Outside Experts reported varying degrees of awareness of the specific contribution of the Program itself. While some respondents could only speak to how the field as a whole had advanced in this area, others were able to speak fairly specifically about how investigators within the Program had contributed to the knowledgebase by confirming linkages between changes in proteomes and the molecular phenotypes of disease.

…those in [Person]'s lab work has really opened up the field and got our viewpoint a little bit broader in this area… Look at [Person]'s numbers, [Person] has enrolled far more people than anticipated by taking the blood of these people and comparing the fragments and, maybe, biomarkers to what you see [Person] knows in the mouse models, there's a lot of advancements out of that.

I'm not familiar enough to cite many specific examples, but one regarding post-surgery inflammation was one that clearly showed a protein profile that we weren't aware of before. So, that's the only specific example I know, but from my interactions with the investigators, it seems that they're getting closer to those sorts of profiles.

…through [Person]'s work, we've been particularly involved with the cardiovascular disease process. So in those areas, there certainly has been some significant moves forward and understanding about protein expression and protein stability. So understanding at the — the physiological level — proteins, there's been a lot of contribution.

A smaller number of Outside Experts questioned whether meaningful progress had been made by the Program in this area.

> *The disease significance of these biomarkers is something we don't really know yet…I think honestly we don't know much more… I think the answer is we develop new technology to answer those questions…we've done a lot of description without a lot of learning about the disease.*

Several Outside Experts commented on the value of the database being generated through the Program. The statements were complimentary, and included comments on the inherent value of the data, and of their value for informing and contributing to the work of other investigators in the field.

> *The data became a very strong tool for discovery…how to annotate the data, analyze the database, and discover normal patterns...so we actually benefitted from their data sets to help our hypothesis in research.*

> *I know the individuals from the other proteomic Centers and their communication is very open. You can go in and talk to them. That's the thing that I find unique about the Centers. Individual grants to scientists…you may be a little bit guarded by saying, "Oh, were not ready to publish that, we don't know that yet"- stuff like that, but with the proteomic Centers, they put the data out there. They say, "here it is, look at it" they make it available.*

> *I think one of the important components of the public funded projects is data dissemination and data sharing. The fact that these data have to be disseminated through various solutions that exist to be publicly available is important. I think that has been one of the major sources of how the dissemination of data, both for NHLBI and a program we have in here, these data need to be published and be available. The raw data as well as the process data, because a lot of proteomics are huge, in terms of the size of the data. There are terabytes of data that are being generated. So the only way that these data could be of use and act upon by the community is to have a sort of point…some sort of place that the data could be available and anybody could download it, no matter where you are in the world. I think NHLBI is an exception and I think they have been doing that, trying to make the data public. I think that's a main source of making data disseminated and should continue.*

Areas the Program Filled Knowledge Gaps in Heart, Lung, and Blood Disease Biology

In general, Outside Experts seemed not to have a sufficiently detailed awareness of the work of the Program to address this question directly. Several respondents commented on the Program's efforts toward identifying and testing new biomarkers, but even these statements tended to be fairly vague. When asked how the Program had contributed to the understanding of underlying mechanisms of heart, lung, and blood biology, Outside Experts declined to comment. One respondent stated that the Program had made meaningful contributions in this area, but did not elaborate.

> *Interviewer: In what ways, if at all, has the Program contributed to our understanding of the underlying mechanisms of heart, love, and blood disease biology? Do you think you would be able to speak to that at all?*
> *Interviewee: No. I could talk to about cancer, but not the NHLBI direction of activities.*

Interviewer: In what ways, if at all, has the Program contributed to our understanding of the underlying mechanisms of heart, love, and blood disease biology? Do you think you would be able to speak to that at all?
Interviewee: I wouldn't want to say what I don't know enough about.

Interviewer: In what ways, if at all, has the Program contributed to our understanding of the underlying mechanisms of heart, love, and blood disease biology? Do you think you would be able to speak to that at all?
Interviewee: Definitely not.

One respondent suggested that this was an overall weakness of the Program, and that the focus on new technologies had eclipsed efforts to improve our understanding of disease biology.

I define innovation as technology, and so the technology has been advancing…my definition of innovative is not necessarily just the new technology, but actually the new cost of understanding. So I would say it's actually made it a step back in that direction…whereas it's a step forward in the development of technologies.

The Extent the Program Investigators Collaborated Within and Outside the Field

In terms of collaboration, Outside Experts provided insight into their own collaborative relationships with investigators from the Program, and in some cases, collaborative relationships between their colleagues and investigators from the Program. Experiences with collaboration varied, though a large proportion of respondents said they had engaged in some form of collaborative relationship, in some cases either newly formed or strengthened by the Program. These respondents spoke highly of these relationships and efforts. Again, there was some lack of awareness of which investigators were involved in the Program.

I've known [Person] off and on for years, probably through [organization]. Very positive [collaboration], we had students that collaborated on building tools and resources. We've taken their data into both the [technology]…it's a very good group to work with, very positive. Stimulating discussion, a lot of times is worth having, but the data has proven to be very high quality over the years as well, which I would say is a big plus for us.

…this is a little bit unique, but [Person] shared data. [Person] has sometimes said, "Hey, would you look for this in your study?" We are glad to do that and it's not only me…I disseminate [Person]'s results whenever I have a chance. I know individuals who have specifically gone and turned around and contacted [Person] and said, "Hey, I heard so and so said this, is it true?" [Person] has shared that sort of information.

Well, our collaboration was through grant mechanisms. So, there were a few investigators from the proteomics consortium that have been involved in some of my grants as consultants or even as investigators on those grants…in fact, we have a current collaboration with [Person]'s new lab at [Center].

A smaller, but non-trivial number of respondents reported that they were not engaged in any collaborative efforts with investigators from the Program. In two of these cases, the Outside Experts cited a perceived lack of fit or overlap between their work and the work of the Program.

I don't think I ever collaborated with anybody, but I haven't looked at the whole list. I know many of these investigators pretty well, but I don't think I collaborated with any of them.

Well, at one point, I came very close to doing a collaboration with [Person], who has one of the Centers. In the end, I don't think the funding was there to do it…but I definitely had talked to [Person] about that. I don't think I have ongoing collaborations with any of the other Centers at the moment.

Well, I have collaborated with some of the [Center] folks doing some stuff, but wouldn't say I used their core facilities…

The Extent Science Informed by the Research was Conducted by the Program

Outside Experts spoke to this question in a variety of ways. One theme that emerged related to the influence of the Program on subsequent science involved the influence of the Program via collaborations, and specifically collaborations within local networks that developed around the Centers.

…because two investigators are in both programs, it's kind of like a bridge, and the techniques and methodologies developed in one go to the other and vice versa…so they first showed this in the NHLBI framework, and now they're applying it in the cancer conditions.

I think one of the nice things about the Center is they've really funded groups of people who have reached out to broader groups…many of them have had impacts more than just their funded labs. They've really enhanced the technical prowess of the whole institution and then has helped a lot of other labs…most of these places that are funded now have cores to help a lot of other people…I think they've had much broader impact than just the actual grant they've funded. That part I'm quite confident of.

The Program, I think, provides a couple of functions for the research community. [Center] provide technical access to the tools, so they develop many, many ways to use cutting-edge technology to drive the quantification and complexity of the analysis and so forth. Technology development has allowed us to makes it feasible to identify thousands of proteins and their change is quantifiable and identify their modifications post-translational modification. So we provide the, what I call, the "biological contents". So, we put in our questions to the [Center], seeking proteomic support to search for the answers to our questions. And with that, we bring the disease contents…so we provide the sort of a context for the proteomic application over time.

Other assessments in this area were much more investigator-centric. Here respondents describe their perceptions of the value of the work of individual investigators in influencing other research in the field.

Going through the website [and commenting on each center]…I know some of the people who are at [Center] and they've been doing nice work. I think this [new Person]'s work is just starting to come along, kind of slow, but I think it's picked up in recent years. I don't know how much this [new Person] has had a huge impact, if you want me to be blunt. I never heard of [new Person], so I don't even know who that is, so I guess I wouldn't say very much there. The same with [new Person]. I know who [new Person], but I can't say much. I haven't felt like there's been huge amounts happening from [new Person]. So, I guess it's been mixed.

I just took a quick look at seeing who the Program has and certainly [Person] has been very big in establishing mainstream technologies in high quality collection of data in proteomics…and [new Person] has done some very nice work in the cardiovascular stuff. [new Person] has done a lot of work and what I like about [Person]'s stuff is it's not the standard sort of mass spectrometry based proteomics…So those are people that definitely made some impact, I think.

Some Outside Experts spoke to the value of the NHLBI Proteomics Centers Program relative to other similar programs or initiatives with which they were familiar. As was true for the investigator-centric assessment, these assessments were mixed. Comments were fairly evenly distributed between those that assessed the Program favorably relative to other work being done, and those that rated other initiatives more highly.

> *…I never had the opportunity to bring in the whole big picture and that's what proteomics does. It brings in an understanding of multiple proteins…it's not a single factor…it's how all these components work. The reason why proteomics is significant in this is that the extracellular matrix, the molecules themselves, are big…[Person]'s work, shows there in a relationship. Some of [Person]'s diagrams about how all those connecting parts work, it blows your mind. But it also is a very significant factor in how to deal with big science and the signaling molecules.*
>
> *I think the work of the Program has helped pushed the field forward. I don't know if I'd say without the Program proteomics would be dead, but I think it's definitely a situation where having a coordinated, funded system like this helps keep those skills moving and alive after some early setbacks.*
>
> *I do think without the Program the field would be much more primitive…they seem to be one of the first groups very interested in sort of a systems analysis…I think that's going to be the future as well. In that sense they have been a little bit ahead of the curve. A lot of other people in institutes talked about it, but most have stepped back and this Program was probably one of the few that stepped up.*
>
> *The [other program] I can speak to because I know it intimately and I think you could ask anybody in the field, that's leading the field, that the NHLBI Program isn't seen as leading the field. I think again my knowledge of this is not deep, but looking over the website there's not the same sort of recognition of the NHLBI Program as the [other program] …it is my guess is this is more focused on individual projects so I wouldn't say this is leading the field… I would know about it if it were leading the field.*
>
> *I think they know they need to partner with clinical folks and to validate the markers they're finding, test them on a larger cohort. I think the groups that know this and are doing this…they've done relative work to [other program] and [other program], any of those groups funded by [funding organization]. I think they've [NHLBI Program] has done better than [other program] which I don't think has done that well at all.*
>
> *So it's not easy for multiple groups to work together even if they want to, but you have to expect that after ten years, they would have co-publications - if they don't, that's a real failure of the Program…If I were to imagine a successful Program, I would imagine after ten years that their stated goals have been achieved. The bar for this project was not very high. They were supposed to identify a proteomic signature. They were not supposed to validate clinically. They didn't have to show that it was meaningful…all they had to do was find a proteomic signature associated with this disease. That's not a very high bar and they haven't done anything. I would be critical of that. That's not a good thing, again, as far as the value of the consortia.*

More specifically, several Outside Experts commented on how the tools and technologies developed through the Program had been shared and applied, both within and beyond heart, lung, and blood research. More generally, a small number of Outside Experts described how the Program had been influential in terms of infusing proteomics into a range of other scientific domains.

I've collaborated with several groups using tools they developed as part of the Program…I've used a technology for sequencing antibodies, I've used antigen microarrays, I've used mass spec technology to study the role of post-translational modifications…

We had some little interaction of using one of their biomarkers…It was more of a technology that they published and we were interested in some of the technology they used and the target they used to work.

Well, I think that the Program has generated a lot of novel methods - that are being widely used, even beyond heart, lung, and blood research. And I'm thinking to a presentation a few years back about post-surgery inflammation and the proteomics response there and I think it's getting very close to clinical patients. But I think probably the main contribution I know of is in the methodology and approach - refining that for other researchers to use.

…one of its major contributions is that more people are now aware of the value of proteomic data. It's now really being used as opposed to just collected.

The Extent the Program has been a Springboard for Subsequent Academic Appointments and Professional Recognition

Outside Experts had varying degrees of knowledge about the specific investigators involved in the Program, and did not comment on the professional advancements or appointments of investigators. However, two Outside Experts specifically praised the Program's work in training young scientists and contributing to the advancement of their careers.

One thing that really impressed me about this group as a whole was the support of young investigators and allowing the young investigators to present their research at these meetings. As a matter of fact, they organized it so that the young investigators could present to our institute director, which is really a big deal.

A number of fellows go through the system and got onto university positions as well as [Person] keeps good contact with industry.

Outside Experts did speak, to some degree, about the professional recognition of some of the lead investigators within the Program, identifying, for example, key academics talks or notable presentations they had given.

I've heard some talks recently from folks from [Person]'s lab. I wasn't sure of the NHLBI connection, but I always thought, "wow, that's just cool"…like the stuff that [Person] doing up there.

The people who are doing well are generally presenting at the major meetings, so I think that's good…for example, they do run workshops at meetings and I think that's good.

[Person] and [new Person] are unique people because they are women. And women scientists certainly have egos. They are not as mad as the men. I'm serious on this, too. How would word that [laughs] or say that…anyway, I find they are much better scientists in terms of their interaction. You know, they go out of their way to be collaborative. They're much more socially involved in discussing their work and presenting it. So, I think they are at the top. I mean, they're presenting their material. Their organizing symposia. They're exchanging things.

Last year, [Person] gave one of the invited presentations at this [specific name] meeting, which is an invitation type of meeting. Prior to that meeting, two years before that, [Person] was one of the prime organizers of this meeting and this is a tough group to break into. There are very strong international symposia builder and [Person] was able to kind of crack the net, you might say, and set up a symposium for the first time. The scientific director commented to [Person]'s symposia that it was one of the best rated ever…so the input there is pretty good.

VII. Critiques and Criticisms

Evaluation of the NHLBI Proteomics Centers Program would not be complete without attention to the criticisms the surfaced from multiple stakeholders. Although the extent to which these criticisms limited the success of the Program is not clear, it is important to highlight key themes that emerged in this area. Overall, four major criticisms were voiced by stakeholders related to either the design of the Program or the overall strategy and direction of the Program.

Design of the Program

Themes emerged related to the design to the Program directly related to funding mechanism and capacity for all three areas of technology development, mechanistic and functional studies, and clinical application.

1. **While many of the center-based leadership indicated they appreciated the flexibility in the design (center-based, collaborative team science) and the funding mechanism (i.e., contracts versus grants, cooperative agreements), others raised questions and concerns related to the funding mechanism.**

   Factors that hindered work? Well, one was just the structuring of a contract versus a grant, where we had milestones and deliverables and couldn't carry funding over. That would have been nicer to have as a grant, but I don't know that would change a whole lot. Also, all the programs budget got a budget reduction because of whatever was available. So you can imagine since you're doing the tools and doing the application optimization, and when you get cut, the thing that gets really harmed is the clinical [application]. –Investigator

   …five years became a learning curve step by step. We were learning what things you could do, what things you could not do…and of the things you could do, how to do them. If the agency had given the principal investigators a two-day workshop ahead of time, I think it would have saved everyone's time. –Investigator

   …there has been scientific progress, I think a lot of it. It's just the question of whether that could have been done under RO1-type funding rather than Center funding. So the big question is always, "Is a Center grant – Is the sum greater than the sum of its parts?" You know, is the total greater than the sum of its parts? And often times the answer is, "No". –Investigator

   That was one of the things with the contract that was a little bit frustrating in that grants, we were allowed to carry funding over with grants, and, often, you can do a no-cost extension for an extra year. But, for this particular contract, it was one of those things that it ends on I think it's August 31st of this year, and once it's done, it's done, so we're all kind of struggling to figure out ways to keep the projects going and to get that last bit of data to be able to put these papers together. -Investigator

   I think the progress was slow, and part of that was because the budgets were very difficult to work with. As you probably know, this was a contract, and so it was made for a very inefficient budgeting environment. Basically, you had to specify how many pipette tips and other things we were using, and if that was different then we ended up having to seek NHLBI Program for approval for re-budgeting. It made it very awkward and stilted…an un-nimble way of being able to react to new opportunities. –Investigator

2. **Many questioned the overall emphasis of the Program in this round of funding, offering critical views on how the Program balanced development with application. Many stakeholders wondered whether the concurrent pursuit of the three areas focal areas - technology development, mechanistic and functional studies, and clinical application - spread resources and capacity too thin.**

   I would say we probably got 80% done of what we wanted to get done. I think the problem that we ran into, was the timeframe to be able to develop the methodologies and then implement them was tight. It took a little longer to develop the methods than we had hoped. I think from the early part of the translational work, there's still a little bit more that I wish we had done. In terms of the human work, I think considering the amount of funding that we had with the contract and the amount of work that we did, I think it was wildly successful. –Investigator

   I feel that this project might have been a little too big, each of these grants. I think that there's sort of a sweet spot and it's somewhere between a couple hundred thousand and a couple million dollars per year. I think that when they're a little too ambitious, you get a lot of good and then you don't get to clean up. I think a more modest amount of something bigger than the usual funding mechanisms, but perhaps not quite as big as what was originally allocated. -Investigator

   So human studies are not inexpensive nor are they easy, and so our portfolio had to be balanced between what we were doing clinically as well as what we were doing with the technological advancement. So we ended up having to manage both of those goals. I think we could have made a lot more progress in the clinical arena if that was the focus of the Program, but we had to both develop technologies as well as apply it to the clinical arena. I think that sort of reflects a little bit of ambiguity at the level of NHLBI as to what the proteomic Centers were. -Investigator

Strategy and Direction of the Program

Themes emerged related to the overall strategy and direction of the Program directly related to cross-center collaborations and uncertainty about the future.

3. **Some expressed disappointment that opportunities for cross-center collaboration were not taken advantage of or systematically pursued. Stakeholders noted that the potential of cross-center collaboration was unrealized, although it was not exactly clear whether it was a weakness of the Program or simply the challenges in collaboration that inhibited collaboration within and across Centers.**

   I think there has been not nearly the amount of interagency or inter-institution cooperation for what's going on. I don't see it. –Investigator

   I think that the Centers didn't develop quite as much a degree of synergy between themselves and they could have done much better with regards to developing joint programs and collaborations. I think that one of my disappointments for the Centers was that we were sort of competitive against one another, rather than working together to solve some significant problems. –Investigator

   It did not function in my view as a consortium. I'm part of another major consortium funded through the [other organization]. It really does function as a consortium - projects together that can be worked on across the Centers that provide value that exceeds the individual laboratories that are funded under the effort. -Investigator

4. **A large part of the negativity and criticism of the Program had to do with the stakeholders' reaction to the discontinuity going forward. It was clear that many of the stakeholders were disappointed to learn that the future of the Program was uncertain. Some reacted strongly to the possibility that the capacity built over the funding period might be lost.**

   I think the main thing that's constraining it is the idea that they don't have any future planned for such a level of sophisticated activity within NHLBI. Therefore, we have to scale back on what we can think of. We've had to recommend new positions for some of our personnel because their careers are important. We can't support them through this, so it means that some have already left and some will be leaving, so it's very disruptive. Which is a major problem because the training of people, particularly in proteomics takes a long time. It's very complicated instrumentation, software. You need to have some experience and if the money flow is interrupted, we cannot keep these people at the Center then of course they go somewhere else, find a new job, go to industry or whatever. –Investigator

   …it's tough with these five-year contracts because by the time you do what you said you were going to do in the first year or two, and then you start creating new things, that takes a couple of years to invent them and then validate them and get them ready to take into a bigger disease. By that time, the contract is already ending, and so that's, I think, why you're seeing it at this time, where our contract is now ending but we're about to take off in these areas. –Investigator

   The proteomics person I have [specific date] has no funds to support them anymore…so I will apparently lose them and in a year or so later, if NHLBI decides they should revise this program then essentially we have to start at the beginning. -Investigator

   And we think we're almost done. We're not even close to being done in the tool development, let [alone] the application. And the question is I cannot do that in an RO1 situation. I'm not allowed to. There's just no other format. So as soon as this contract ends…, essentially tool development here except for baby steps I think is going to end. And that I think is a failure of NIH. I mean I'm going to be blunt. -Investigator

VIII. Conclusion

The overarching goal of the Program is to help facilitate a better understanding of the underlying mechanisms in heart, lung, and blood diseases which could contribute to more effective diagnoses, risk stratification, intervention, and prevention. Towards this goal, each of the NHLBI Proteomic Centers encompasses three components to make an interactive team that includes proteomic technology development; mechanistic and functional studies; and application to clinical questions.

The Program was observed to be successful in terms of the subjective and objective measures used to assess the outcomes. Subjectively, program stakeholders viewed the Program as successful on many fronts including creating a stronger network of collaboration. Despite criticisms about certain components of the Program, there was widespread agreement that this initiative was of value to the field of proteomics.

Several celebrated and highly visible technological advancements were attributed to the Program by stakeholders and many felt that the Program had made significant headway in advancing knowledge about the physiologic functioning of the proteome. It was clear throughout this evaluation that from those commenting on the Program, more work was to be done in the area of clinical application. A consistent theme that emerged from stakeholders was the Program simply ran out of time to achieve the clinical goals. Nevertheless, several examples of where the scientific work of the Program was being applied in ongoing clinical studies were cited.

Stakeholders recognized the Program's expanded capacity to engage in proteomics research that extended well beyond the seven Centers. Acknowledgement of collaborations, partnerships, promising early-career professionals, leveraged funding, physical and virtual laboratories, and data clearinghouses describe the establishment of an infrastructure capable of contributing to the advancement of proteomic science. Although much more work is to be done, it is in the estimation of the Program stakeholders that the Program was successful in advancing proteomic science by establishing a new set of tools, knowledge, and capacity to continue the study of innovative proteomic approaches in heart, lung, and blood disease research.

Objectively, through this evaluation, we found evidence of a highly recognized and scientifically informative corpus of scientific output. Subsequence science was being informed by the work produced by the Centers, as other scientists from a variety of different disciplines acknowledged the Program's scientific output in their own work. This collective body of scientific work, as represented by the substantive volume of peer-reviewed papers, was performing at a high level compared to what was expected based on where and when the work was published. Scientific papers produced by the Program were found in many of the top journals in the field.

Furthermore, an extensive network of collaborators contributed to the production of highly regarded scientific outputs. The growth in this network of collaborators over time suggests the Program was very successful in expanding the scientific productivity well beyond the intuitional boundaries established by the Program funding. The Program's commitment to the development of scientists in the field of proteomics was evident in the honors, awards, advancements, and new funding accumulated by trainees and key personnel active in the Program. The wide dissemination and sharing of program information and results was confirmed in the global presence of the Program, both physically and virtually.

Ultimately, we conclude based on the results of this evaluation that the NHLBI Proteomics Centers Program met most of its design expectations. Given that the vison of the NHLBI's Proteomics Center Program was to better understand heart, lung, and blood disease biology, the Program (as a center-based

collaborative research initiative) clearly promoted the application of proteomic technologies to gain an understanding of physiologic pathways for defined clinical questions.

The Program provided funding for some of the top scientists in the world to focus on key scientific problems with a modicum of flexibility in the inquiry, and the successes observed in this evaluation are reflective of this effort. In that sense, this evaluation yielded the kind of information that documents success in the context of a large scale scientific. Although we heard criticisms of the Program, particularly with regard to the funding mechanism, limits to cross-center collaboration, scientific emphasis and balance, and discontinuation of the Program, none of these criticisms were substantial enough to argue the Program was not successful in meeting is programmatic objectives.

It should be noted that the Outside Experts present opinions based on their experience with and knowledge about the Program. They represent a perspective that is a function of the group and their relationship to the Program. While Outside Experts were invited to participate based on a random selection from a list provided to us by NHLBI CORs, the final cohort was based on their willingness to agree to an interview. We did not hear any information from the Outside Experts that contrast the subjective or objective results in a way that would argue against the overall success of the Program.

The design of this evaluation does not permit us to make a claim about causality. In other words, it is not possible for us to demonstrate unequivocally that the Program directly led to the outcomes observed or that the scientific success recognized by the researchers over the evaluation period was due to the Program and without it, these successes would not happen. However, we are able to make a plausible connection of the Program design, the resources committed, and the contribution to the capacity and scientific successes manifest. Indeed, an argument could be made that the center-based, collaborative research model may accelerate progress toward resolving complex societal and scientific problems. Upon this model we make the assumption that the NHLBI Proteomics Centers Program expedited scientific inquiry and discovery, and reinforced the belief that team-based research integrating the strengths of multiple disciplines. In this regard, we have made a case that, based on the Programs espoused theory of change, it is reasonable to conclude the results of the Program were attributable to the interactive elements (i.e., people, resources, activities, motivations, etc.) of the Program.

IX. Appendices

Appendix A. Expert Advisory Panel

Appendix B. Evaluation Data Source Matrix

Appendix C. Key Informant Interviews and Qualitative Analysis

Appendix D. Data Abstraction/Scientific Literature Review, Bibliometrics, Altmetrics, and Publication Analysis

Appendix A: Expert Advisory Panel

Proteomics Experts

Dr. Ian Blair, PhD
Director, Proteomics and Systems Biology Faculty
University of Pennsylvania
Philadelphia, PA
ianlblair@exchange.upenn.edu

Dr. Blair joined the University of Pennsylvania in 1997 as the A.N. Richards Professor of Pharmacology. In 2002, Dr. Blair was named the Vice-Chair of the Department of Pharmacology and in 2003 he established a mass spectrometry-based Proteomics and Systems Biology Facility at the University of Pennsylvania. Dr. Blair is an internationally recognized expert in the utility of mass spectrometric methods and has published over 290 manuscripts.

Dr. Blair received his Ph.D. in Organic Chemistry in 1971 from Imperial College of Science and Technology in London and held research fellowships at the Australian National University in Canberra and at Adelaide University in Australia. Dr. Blair has been recognized with honors and awards for his distinguished achievements in mass spectrometry.

Dr. Robert L. Moritz, PhD
Director, Proteomics Research Laboratory
Institute for Systems Biology
Seattle, WA
rmoritz@systems.org

Dr. Mortiz joined the Institute for Systems Biology (ISB) in 2008 as Director of Proteomics and Associate Professor where he is the leader of an initiative to establish a Human Multiple Reaction Monitoring (MRM); a resource to provide scientists ability to conduct quantitative analysis on all human proteins. Previous tenure included appointment at the Ludwig Institute for Cancer Research in Melbourne, Australia for over 25 years where he established a collaborative bioinformatics center. Dr. Moritz is recognized for his expertise in design and implementation of multiple technologies currently used in proteomics laboratories globally.

Dr. Moritz received his PhD from the University of Melbourne. Dr. Moritz has been recognized with honors and awards for his professional efforts and achievements in bioinformatics and targeted proteomics.

Dr. Russell Bowler, MD, PhD
Professor of Medicine
Division of Pulmonary, Critical Care and Sleep
National Jewish Health
Denver, CO
BowlerR@NJHealth.org

Dr. Bowler is a Professor of Medicine in the Department of Medicine, Division of Pulmonary, Critical Care and Sleep Medicine at National Jewish Health and Principle Investigator of Bowler Laboratory. Dr. Bowler's current projects in the Bowler Laboratory include GWAS, Genome, Metabolome and proteome

profiles of emphysema and airway disease, cigarette smoke induces endogenous oxidant injury and textural approach to quantification of diffuse lung disease on CT.

Dr. Bowler earned a B.S. in Mathematical and Computational Sciences from Stanford University, a M.D. from the University of California at San Francisco, and a PhD in Cell and Developmental Biology from the University of Colorado. He completed fellowships in internal medicine residency at University of California at San Francisco and pulmonary and critical care fellowship from Colorado University.

Evaluation Experts

Brian Zuckerman, PhD
Technology Policy and Assessment Center
School of Industrial and Systems Engineering
Washington DC
bzuckerm@ida.org

Dr. Zuckerman is a member of the Research Staff at the Science and Technology Policy Institute (STPI) in Washington DC. Dr. Zuckerman's work concentrates on Federal research and development program performance and agency-wide research portfolios with focus on biomedical research. He is recognized for his expertise in areas of program evaluation and scientometrics. Dr. Zuckerman has analyzed Federal research and development data systems, and statistical data collection programs. Prior to joining STPI, he was a principle a C-STPS, LLC and at the Center for Science and Technology Policy of Abt Associates Inc.

Dr. Zuckerman holds a BA in Chemistry from Harvard, and a PhD in Technology, Management and Policy from the Massachusetts Institute of Technology. He is a co-chair of the Research, Technology and Development Topical Interest Group of the American Evaluation Association.

Gretchen Jordan, PhD
360 Innovation LLC
PO Box 51404
1135 Shell Ave.
Pacific Grove, CA 93950
gretchen.jordan@comcast.net

Dr. Jordan is an independent consultant specializing in a systems view of innovation and program and evaluation design. Dr. Jordan was previously a Principle Member of the Technical Staff with Sandia National Laboratories working with the U.S. Department of Energy (DOE) Office of Energy Efficiency and Renewable Energy, and the DOE Office of Science, on evaluation and performance measurement and innovative methods of assessing the effectiveness of research organizations.

Dr. Jordan has a Ph.D. in Economics. She is a Fellow of the of the American Association for the Advancement of Science, and co-founder and co-chair of the American Evaluation Association's Topical Interest Group on Research, Technology, and Development Evaluation.

Dr. Juan Rodgers, PhD
Professor
School of Public Policy
jdrogers@gatech.edu

Dr. Rogers is an Associate Professor of Public Policy and Director of the Research Value Mapping Program at the School of Public Policy, Georgia Institute of Technology. His current research interests include modeling the R&D process, assessment of R&D impacts with specialized focus in the formation of scientific and technical human capital, technology transfer, R&D policy and evaluation, the interaction of social and technical factors in the development of information technology, contextual factors of scientific and technical creativity, and information technology policy. He is recognized internationally as a consultant on science, technology and innovation policy and holds many publications in the field.

Dr. Rogers earned a PhD in Science and Technology Studies at Virginia Polytechnic Institute and State University. His professional efforts have been honored and distinguished with multiple awards and grants.

Appendix B: Evaluation Data Source Matrix

	Key Informant Interviews					Publication Analyses			
Data Source	Internal	External	Annual Report	Biosketches /CVs	Web analytic	Bibliometric Data	Altmetric Data	Co-author	Scientific Aims
Evaluation Questions/Domain									
1. To what extent has the program developed and shared new tools and	**X**	**X**	**X**		**X**		**X**		
2. What outputs of the program have matured enough to be of clinical utility? To what extent has the program resulted in inventions, patent applications, spin-off companies, and	**X**	**X**	**X**		**X**	**X**			
3. To what extent has the research contributed to the creation and integration of a knowledgebase linking changes in proteomes with	**X**	**X**			**X**	**X**			
4. In what areas has the program filled existing knowledge gaps in heart, lung, and blood disease	**X**	**X**	**X**			**X**	**X**		**X**
5. To what extent have the program investigators been collaborative within and	**X**	**X**	**X**			**X**		**X**	
6. To what extent was subsequent science informed by the research	**X**	**X**				**X**	**X**		
7. To what extent has the program served as a springboard for subsequent academic appointments and	**X**	**X**	**X**	**X**					

Appendix C. Key Informant Interviews and Qualitative Analysis

C.0. Qualitative Interviews Procedures and Protocols

Qualitative Data on the NHLBI Proteomics Centers Program were collected with implementation of in-depth semi-structured interviews to capture information about the purpose and impact of the Program, the structure and organization of the Program, and the roles and activities of individuals involved with the program. Interview guides were developed and used to facilitate the interview process with individuals in four targeted groups, which included: (1) principal investigators (PIs) currently involved in the Program; (2) trainees who are currently or were previously involved in the Program; (3) Outside Experts in the field of proteomics who are not directly involved with the Program, and; (4) representatives from NIH, NHLBI, and NCI (referred to collectively as NIH) who have knowledge of or who have participated with the Program in some capacity. Individuals from these respective groups represented unique perspectives that informed successes and limitations of the program. The number of interviews was broken down by group as follows: 27 PIs (including Center Directors, Co-PIs, and Investigators), 20 trainees (including both current and former), 9 Outside Experts, and 9 NIH **(Table 11 below)**. In order to reach this goal, interviews were conducted at the NHLBI PI meeting, by phone, or by Skype.

Table 11. Interview Respondents by Group

Respondent Group	Target # of Interviews	Completed # of Interviews
Principal Investigators (PIs)		
Directors	7	7
Co-PIs	14	12
Investigators	6	6
Trainees		
Current	10	9
Former	10	9
Outside Experts	9	9
NIH	9	8
Total	**65**	**60**

C.1. Participation

Two sets of interviews were conducted with respondents that were drawn from the four respective groups (as described above). The interview data collection required approval from the Office of Business and Management (OMB). However, given the nature of the interviews, IRB approval was not required though respondents were reminded of their voluntary participation. OMB approval for the interview data collection was approved in May 2015. The first set of interviews were completed during the period of January 20 – February 17, 2015. The second or "full" set of interviews were completed during the period of June 14 – October, 7 2015. The interviews were designed to take approximately 30-45 minutes, though a small number of interviews exceeded this timeframe.

C.1.1. Pilot Interviews: Respondents

Respondents were drawn from the four groups (as described above) with knowledge of NHLBI Proteomics Centers Program or of the field of proteomics more generally. A list of potential respondents from each of these four groups was compiled with the assistance of partners at NIH/NHLBI and the Center Directors. Respondents were recruited for the pilot interviews using stratified random sampling without replacement. Within each of the four groups (strata), potential respondents were randomly selected. A total of fourteen pilot interviews were conducted with three PIs, three trainees, three Outside Experts, and five representatives from NHLBI/NIH/NCI. Among the PIs and trainees were representatives from five of the seven Centers.

C.1.2. Full Interviews: Respondents

As with the pilot interviews, respondents in the full set of interviews were drawn from same respective groups as described above. Recruitment for the key informant interviews was influenced by (1) the restrictions on the total number of interviews per stakeholder group, (2) the expectation, that the majority of interviews were conducted in-person at the NHLBI Proteomics Centers Program PI meeting in June 2015 as described in the task order, and (3) there was significant variation in the number of individuals in each role within the Centers. In order to gain as comprehensive of an understanding of the Program from the perspective of key stakeholders, and to avoid over representing any one group, a mixed purposeful sampling strategy was employed that allowed recruitment goals within these boundaries to be met.[17]

In addition to the overall targets for each stakeholder group, targets were set for meaningful subgroups within the PI group, including Center Directors, Co-PIs, and other investigators within the Centers. Interviews were conducted with all seven Center Directors and up to three Co-PIs from each Center (due to a wide variation in the number of Co-PIs per Center and concerns about overrepresentation from a single Center). A total of six other investigators from across the Centers were also interviewed.

C.3. Recruitment

For the pilot interview, these individuals were identified from a list described above in ***C.1.1. Pilot Interview: Participation***. Once identified, the respondent was contacted via phone and/or email by the evaluation team, and asked to participate in an interview. Recruitment continued until the target number of respondents for each group was met. The pilot interviews were conducted between January 2015 and February 2015. A total of 14 pilot interviews were conducted and included three PIs, three trainees, three Outside Experts, and five representatives from NIH. Among the PIs and trainees, five of the seven Centers were represented.

For the full set of interviews, under the terms of the contract, to maximize the number of in-person interviews conducted, respondents were recruited at the NHLBI PI meeting. Individuals were identified from a list described above in ***C.1.1. Pilot Interview: Participation***. Respondents were identified based on a list of attendees provided to CSI by NHLBI. Using the mixed purposeful sampling strategy, potential respondents were sampled from three participating stakeholder groups: PIs, trainees, and NIH. Attendance at the meeting was lower than anticipated. However, the evaluation team completed 26 of the 65 interviews (40%) onsite at the meeting.

Following the NHLBI PI meeting, respondents were recruited for the remaining interviews using stratified random sampling without replacement. For internal Center staff, potential respondents were generated based on a full list of staff including trainees and former trainees. Potential respondents from NIH and Outside Experts were compiled with the assistance of partners at NIH and the Center Directors. Potential respondents were contacted via phone and/or email by the evaluation team, and asked to participate. Interviews were conducted through early October 2015 at which point we had completed a total of 60 interviews. The interview component of the evaluation was closed with 92% of the projected interviews complete (60 of the 65) **(Table 11 above)**. This is due in part to the availability of the respondents and saturation of interview data.

C.4. Interview Guides

The interview questions were framed by the evaluation logic model, evaluation questions, and with input from the NHLBI Project CORs and Expert Advisory Panel. The interviews were designed to yield appropriate information for addressing the evaluation questions by maximizing a "structured flexibility" approach across respondent groups. In total, four separate interview guides were developed for each of

[17] Patton, MQ (2002). *Qualitative evaluation and research methods*. 2nd ed. Thousand Oaks, CA: Sage.

the four groups, with a subset of questions included in all guides. The interview guides were designed in a way that enabled individuals to provide information shaped by their experiences, in the context of the overarching evaluation questions. Feedback and analysis from the pilot interviews was essential to help structure and refine the four interview guides for the full interview phase of the project. The full interview guides for each respondent category can be found in ***C.8. Full Interview Guides***.

C.5. Data Collection

Data collection and analysis procedures were synchronous and iterative for both sets of interviews in keeping with the emergent nature of qualitative methodologies. Data collection included a series of steps: receiving consent from the interviewee to participate, recording the interview, internal memoing and debriefing among members of the evaluation team, and interview transcription.

C.5.1. Respondent Consent

As mentioned above, the evaluation had OMB clearance and did not need IRB approval. Even though a signature was not needed from respondents, off the record, the evaluation team asked each respondent if they agreed to the conditions of the interview, understood that this interview was confidential, understood that their participation was voluntary, and asked to for permission to audio record the interview. Once the interviewer gained a verbal consent and answered any questions from the interviewee, the interview was then conducted.

C.5.2. Memo

After each interview, the interviewer documented methodological and analytical reflections on the interview through a standardized memo process. The interviewer then participated in a process of peer debriefing with another member of the evaluation team. In addition to documenting key reflections on the content of the interview, this allowed the evaluation team to identify challenges and make minor adjustments (such as adding probes) to improve the flow and function of the interview guides, and to ensure that key concepts that emerged during the process could be explored in subsequent interviews.

C.5.3. Transcription

After the interview was conducted and the memo process was complete, the audio recording was transcribed by a third-party service for analysis. This process was conducted virtually where the audio file was emailed to the transcription service. Once the audio file was transcribed, the interview document was sent back to the evaluation team for analysis. A detailed record was maintained throughout this process of when audio files were distributed for transcription and when the transcription was complete.

C.6. Quality Assurance

Quality assurance was maintained through a two-part approach. First, the peer debriefing process served as a quality assurance measure by identifying technical or procedural issues that needed to be addressed after each interview, allowing the evaluation team to make real-time adjustments to improve the process. Second, a subset of interviews were audited by the Senior Evaluator, and assessed using a standard quality review checklist, found in ***C.9. Quality Control Checklist***, designed to ensure key elements of the interview process were consistently and appropriately implemented. Audits for the interview data were implemented in two phases. This process was used for both sets of interviews.

The audits followed the completion of a specific checklist developed by the Senior Evaluator to review the interview output. In order to gauge the comprehensiveness of the interview in relation to the guide, a detailed checklist indicating presence (or absence) of specific contextual information and content was completed with every third interview during the initial stages of the process (first 30 interviews completed), for each respective interviewer, separately. Interview output that did not meet specific elements on the checklist were flagged and submitted for review by the internal evaluation team for

follow up to correct information or fill in missing data. The review process enabled the interviewers to make adjustments during the early stages of the process of interviewing. Therefore, completion of the checklist was done at wider intervals as the team proceeded with interviews, specifically every sixth interview.

The second set of audits for the qualitative interview data was conducted by the Senior Evaluator to ensure the quality of transcription by comparing the transcribed interviews against the actual recording. A random sample of interviews were selected from the final set (n = 8) of transcriptions and were reviewed for accuracy. Findings were reported and reviewed during internal weekly evaluation meetings, which enabled the team to catch and correct problems as necessary.

C.7. Analysis

The primary purpose of the key informant interview process was to understand the Program as comprehensively as possible from multiple perspectives. Following single embedded case study methodology, which allows for data to be analyzed at both the program and, where appropriate, the stakeholder level, interview data were analyzed both deductively and inductively.[18] The qualitative lead initiated the coding and data reduction process by developing a codebook based on a priori themes identified from the program logic model. Then, beginning with the transcripts of three randomly selected interviews, open coding was used to identify key themes that emerged from the data. The codebook was adjusted to reflect the themes identified through the inductive analysis. The codebook can be found in ***C.10. Analysis Codebook***.

The codebook was reviewed by the evaluation team, and the remaining interviews were coded by the qualitative task lead as the transcripts were completed. Intercoder agreement was addressed by having two trained members of the team independently code, and then compare, a set of interviews with those coded transcripts completed by the qualitative task leader. The coders discussed and reached consensus on points of disagreement or uncertainty. Throughout the process, proposed adjustments to the codebook were discussed and, if appropriate, implemented. The final codebook consisted of 35 themes, plus subthemes.

Once the codebook was developed, it was inputted into OSR NVivo Data Analysis Software (referred to as NVivo) to analyze all the interviews. In total, roughly 40 hours of interviewing was conducted, producing 962 typewritten pages of transcripts. Using NVivo software, the evaluation team developed a simulated matrix model from the coded material. This produced data on (1) the total number of references associated with each code, and (2) the number of times there was overlap between codes. By examining the overlap between codes, we are able to see, conceptually, the number of times adjacent codes were related, conceptually, in the interview data. Using the logic model as the foundation for this evaluation, the coded matrices were then overlapped onto the logic model to visually display areas of overlap in the data found in ***C.11. Codebook Matrix***. The matrix analysis displays the results with the values inside the logic model nodes indicating the total number of references for that node, and values along the connections indicating the number of references in which the connected nodes were jointly referenced.

[18] Yin, RK (2014). *Case study research design and methods*. 5th Ed. Thousand Oaks, CA: Sage.

C.8. Full Interview Guides

NHLBI Proteomic Centers Program Study
Key Informant Interview Guide – Investigators

Interview Information

Interviewee:		
Title:		
Interview Format	□ In-person □ Phone □ Skype/video chat	
Phone/Email:		
Video chat ID:	Skype:	GooglePlus:
Date:	Start time:	End time:
Interviewer:		

Introduction

Opening Script:

- Thank you for taking the time to participate in this study.
- My name is _____, and I am with Concept Systems, Inc. We are an evaluation firm that has been contracted to conduct an evaluation of the NHLBI Proteomics Centers Program (which I will refer to as "the Program" from now on).
- Goals of the evaluation are to better understand the contributions of the Program, and the experiences of Program staff and affiliates.
- Looking at the Program as a whole. Findings will not be used to compare the individual Centers.
- Interview should take about 45 minutes.
- This evaluation has been approved by the OMB. Not research, so IRB approval not required. However, your participation is voluntary – can skip questions or stop.
- The evaluation only covers the past five years of the Program, from 2010 to 2015. If you were involved in previous iterations of this program, please make sure that your responses reflect your experience only over the past five years.
- Info you provide will not be disclosed to anyone but the researchers conducting the study, except otherwise required by law.

With that said, are you willing to continue with the interview? [YES/NO].
And is it okay if I audio record the interview? [YES/NO]
Do you have any questions before we begin? Let's get started.

Background Questions

Script: First, I want to gather some basic information about your role and position.

1. Please verify the following background information:

Title:	
Center/Organization:	

2. During what years have you worked at the Center?
    a. <u>If before 2010, prompt</u>: In responding to my questions, please remember that we are focused only on the past 5 years.
3. Did your position or title change during that time?
4. Very briefly, how would you describe your current role at the Center?

Section A: Scientific Advancements & Influence of the Program

Script: I would like to begin by asking you some questions about the scientific advancements made through the Program over the past five years, and the influence of the Program in the fields of proteomics and heart, lung, and blood research.

5. Thinking back over the past five years, what would you say are the top two or three major discoveries or advancements to come out of your Center?
    a. How, if at all, would you say these discoveries have influenced proteomics and heart, lung, and blood research?
6. How, if at all, has the Program increased the capacity of the field as a whole to engage in innovative proteomics and heart, lung, and blood research?
    a. What role, if any, has collaboration played in influencing the capacity of the field?
7. Thinking about the connection between proteomes and the molecular phenotypes of disease, what do we know that we did not know 5 years ago, based on the work of your Center?
    a. In what ways, if at all, has your Center contributed to our understanding of the underlying mechanisms of heart, lung, and blood disease biology?

Script: The next questions ask specifically about clinical and translational advancements. I'll start by asking about the work of your Center, then move to the Program, then move on to think of the field as a whole.

8. Broadly speaking, what were the clinical or translational goals that your Center set out to achieve?
9. How would you describe the progress that your Center made in achieving its clinical and translational goals? (Probe: Where there any specific discoveries, knowledge advances, or tools emerged from the work of your Center that contributed to achieving these goals?)
    a. How would you describe the next steps needed to fully achieve these goals?
    b. In your opinion, what role, if any, did the 3-part interactive team structure of the Program play in making progress toward these goals?

Script: Now I'm going to ask a similar question, but in regard to the Program as a whole.

10. In your opinion, how has the Program, advanced the field in terms of clinical outcomes?
11. Now, thinking about the field as a whole, how would you describe the next steps needed in terms of clinical applications of proteomic research? In other words, where are we now, as this five-year phase of the Program is wrapping up?

Section B: Evolution of the Centers and Program (Investigators only)

Script: Thank you. Now I would like to ask you a few questions about factors that have influenced the work of your Center.

12. In the past five years, what, if any, changes have your Center undergone in terms of direction? In other words, how, if at all, has the overall focus or purpose of the Center evolved?
    a. If there was a change in focus, how was that change implemented?
13. What factors, if any, have constrained or hindered the work of your Center?
14. In what ways, if at all, has the work of your Center been influenced by discoveries or advancements made elsewhere in the fields of proteomics and/or heart, lung, and blood research?

Section C: Application & Dissemination of New Knowledge and Tools

Script: Great. Now I am going to ask you a series of questions related to the application and dissemination of knowledge and tools developed through the Program.

15. Within your Center, how is new knowledge shared and integrated across researchers and labs?
    a. How is this process similar to or different from activities facilitated through other grants or contracts?

16. In what ways, if at all, did researchers at your Center access new knowledge, methods, or tools created through the other Centers?
    a. In what ways, if at all, was this information integrated into the work and research of your Center?
17. In what ways, if at all, were knowledge and tools developed at your Center made available to the broader proteomics and NHLBI research communities?
    a. Please describe the process and timeframe through which tools and prototypes were made available.
18. In what ways, if at all, has the sharing of knowledge and tools influenced the overall process of discovery for the Program?

Section D: Collaboration

Script: Now I would like to ask you about the role that collaboration played in your work and the work of your Center.

19. How would you characterize the collaborations that stemmed from your work at the Center? For example, with whom did you tend to collaborate, what was the nature of the collaborations, and how would you describe the quality of the collaborations.
    a. In what ways, if at all, did these collaborations impact your work at the Center?
    b. How, if at all, has the focus of the Program shaped decisions regarding collaboration within your Center?
20. What, if any, achievements or accomplishments of your Center might not have been successful without collaboration, in your opinion?
    a. How would you describe the role of collaboration in these instances?

Script: Now I would like to ask a question about the more formal collaborative relationships that may have developed with other Centers or researchers outside of the Program.

21. In your work at the Center, what, if any, formal collaborations did you participate in that involved working collectively with researchers outside of the Center to achieve a common goal?
    a. How would you describe the nature of these collaborations?
    b. In your opinion, how valuable were those collaborations in terms of accomplishing your goals or the goals of the Center? For example, would you say that they were instrumental, generally productive, or generally unproductive?

Section E: Recognition & Advancement

Script: My final question focuses on the recognition and professional advancement of researchers involved in the Program.

22. In what ways, if at all, has your work at the Center contributed to the advancement of your career?
23. That is the last of my formal questions. Before we close, is there anything about the Centers or the Program that you feel is important for us to know, that I didn't ask about?

Closing Remarks

Script: Thank you again for taking time out of your schedule to participate in this study. We will be using the data we collect through these interviews to inform our evaluation of the Program. If, after revisiting our interview, I have questions for clarification, is it okay if I contact you again? [YES/NO]
Great. Thank you again for your time. Please feel free to contact us if you have any follow-up questions or comments.

Interviewer notes

Methodological comments
• How did the process go? What worked well? What didn't? • Observations on the questions/guide: (Redundancy; Flow; Specificity/generality of the questions) • Other
Analytical comments
• Thoughts/observations on the content of the interview • Themes or connection • Demeanor of interviewee • Key new information • Other

NHLBI Proteomic Centers Program Study
Key Informant Interview Guide – Trainees

Interview Information

Interviewee:		
Title:		
Interview Format	□ In-person □ Phone □ Skype/video chat	
Phone/Email:		
Video chat ID:	Skype:	GooglePlus:
Date:	Start time:	End time:
Interviewer:		

Introduction

Opening Script:

- Thank you for taking the time to participate in this study.
- My name is _____, and I am with Concept Systems, Inc. We are an evaluation firm that has been contracted to conduct an evaluation of the NHLBI Proteomics Centers Program (which I will refer to as "the Program" from now on).
- Goals of the evaluation are to better understand the contributions of the Program, and the experiences of Program staff and affiliates.
- Looking at the Program as a whole. Findings will not be used to compare the individual Centers.
- Interview should take about 45 minutes.
- This evaluation has been approved by the OMB. Not research, so IRB approval not required. However, your participation is voluntary – can skip questions or stop.
- The evaluation only covers the past five years of the Program, from 2010 to 2015. If you were involved in previous iterations of this program, please make sure that your responses reflect your experience only over the past five years.
- Info you provide will not be disclosed to anyone but the researchers conducting the study, except otherwise required by law.

With that said, are you willing to continue with the interview? [YES/NO].
And is it okay if I audio record the interview? [YES/NO]
Do you have any questions before we begin? Let's get started.

Background Questions

Script: First, I want to gather some basic information about your role and position.

1. Please verify the following background information:

Title:	
Center/Organization:	

2. During what years have you worked at the Center?
    a. <u>If before 2010, prompt</u>: In responding to my questions, please remember that we are focused only on the past 5 years.
3. Did your position or title change during that time?
4. Very briefly, how would you describe your current role at the Center?

Section A: Scientific Advancements & Influence of the Program

Script: I would like to begin by asking you some questions about the scientific advancements made through the Program over the past five years, and the influence of the Program in the fields of proteomics and heart, lung, and blood research. I understand that there are multiple levels of expertise and experience within the Center, but we are really interested in your perspective as a trainee.

5. Thinking back over the past five years, what would you say are the top two or three major discoveries or advancements to come out of your Center?
    a. How, if at all, would you say these discoveries have influenced proteomics research and heart, lung, and blood research?
6. Broadly speaking, what were the goals that your Center set out to achieve related to research applicable to clinical or translational questions?
7. What, if any, specific discoveries, knowledge advances, or tools emerged from the work of your lab or group that contributed to achieving these goals?

Section B: Application & Dissemination of New Knowledge and Tools

Script: Great. Now I am going to ask you about the application and dissemination of knowledge and tools developed through the Program.

8. In what ways, if at all, did you and your colleagues access new knowledge, methods, or tools created through the other Centers?
    a. In what ways, if at all, was this information integrated into your work and research?

Section C: Collaboration

Script: Now I would like to ask you about the role that collaboration played in your work and the work of your Center.

9. In your role with the Center, in what ways, if any, have you been able to collaborate with other researchers outside of your lab or group?
    a. With other researchers in the field?
10. How would you describe the value of these collaborations to your training and professional development?
    a. To the work of your lab or Center?
11. In what ways, if any, do you intend to leverage or build on these collaborations going forward?

Script: My next question is about more formal collaborations, specifically collaborations that involved working collectively with researchers outside of your lab or group to achieve a common goal.

12. In your work at the Center, in what ways, if any, did you engage in formal collaborations with other researchers outside of your lab or group? (Note: Here we are referring to collaborations that involved working collectively with other researchers to achieve a common goal.)
    a. With researchers outside of your Center?
    b. How would you describe the nature of these collaborations?
    c. How would you describe the value of these collaborations for your training or professional development?
    d. How would you describe the value of those collaborations for the work of your lab or group? For example, would you say that they were instrumental, generally productive, or generally unproductive?
13. Which of the following three terms best describes your training before joining the Center: Would you say your training was primarily in (1) proteomics/computation, (2) basic science/molecular studies, or (3) clinical studies?
    a. How, if at all, has your work at the Center influenced your interest in incorporating elements of the other two areas into your work going forward?

b. Probe: For example, if your training was clinically focused, has your work at the Center made you consider incorporating computation or basic science into your work going forward? Or, has your work at the Center influenced you toward a more interdisciplinary approach?

Section D: Recognition & Advancement

Script: My final questions focus on the recognition and professional advancement of researchers involved in the Program.

14. I am going to ask you to consider the work and standing of the lead investigators involved in the Program, relative to other leading investigators in the field. How would you characterize the leading investigators who are involved with the Program?
    a. For example, would you say they are on par with other investigators, exceptional in the field, or lagging behind other senior investigators in the field?
15. In what ways, if at all, has your work at the Center contributed to the advancement of your career?
16. How, if at all, do you expect to apply the skills and knowledge you gained at the Center in your career moving forward?
    a. Do you intend to continue working in the field of proteomics? If so, in what capacity? If not, why?
17. That is the last of my formal questions. Before we close, is there anything about the Centers or the Program that you feel is important for us to know, that I didn't ask about?

Closing Remarks

Script: FOR FORMER TRAINEES ONLY: Before we close, I would like to ask you about your level of familiarity with the Program.

18. At this point, would you say that you have: a) considerable familiarity, b) some familiarity, or c) little or no familiarity with the Program? [Circle one]

Script: Thank you again for taking time out of your schedule to participate in this study. We will be using the data we collect through these interviews to inform our evaluation of the Program. If, after revisiting our interview, I have questions for clarification, is it okay if I contact you again? [YES/NO]
Great. Thank you again for your time. Please feel free to contact us if you have any follow-up questions or comments.

Interviewer notes

Methodological comments
• How did the process go? What worked well? What didn't? • Observations on the questions/guide: (Redundancy; Flow; Specificity/generality of the questions) • Other
Analytical comments
• Thoughts/observations on the content of the interview • Themes or connection • Demeanor of interviewee • Key new information • Other

NHLBI Proteomic Centers Program Study
Key Informant Interview Guide – Outside Experts

Interview Information

Interviewee:		
Title:		
Interview Format	□ In-person □ Phone □ Skype/video chat	
Phone/Email:		
Video chat ID:	Skype:	GooglePlus:
Date:	Start time:	End time:
Interviewer:		

Introduction

Opening Script:

- Thank you for taking the time to participate in this study.
- My name is _____, and I am with Concept Systems, Inc. We are an evaluation firm that has been contracted to conduct an evaluation of the NHLBI Proteomics Centers Program (which I will refer to as "the Program" from now on).
- Goals of the evaluation are to better understand the contributions of the Program, and the experiences of Program staff and affiliates.
- Looking at the Program as a whole. Findings will not be used to compare the individual Centers.
- Interview should take about 45 minutes.
- This evaluation has been approved by the OMB. Not research, so IRB approval not required. However, your participation is voluntary – can skip questions or stop.
- The evaluation only covers the past five years of the Program, from 2010 to 2015. If you were involved in previous iterations of this program, please make sure that your responses reflect your experience only over the past five years.
- Info you provide will not be disclosed to anyone but the researchers conducting the study, except otherwise required by law.

With that said, are you willing to continue with the interview? [YES/NO].
And is it okay if I audio record the interview? [YES/NO]
Do you have any questions before we begin? Let's get started.

Background Questions

Script: First, I want to gather some basic information about your role and position.

1. Please verify the following background information:

Title:	
Center/Organization:	

2. In what capacity have you worked with the Program, or how are you familiar with the Program?
3. In responding to my questions, please remember that we are focused only on the past 5 years of the Program. Has your role with the Program changed during that time?
4. Very briefly, how would you describe your current role?

Section A: Scientific Advancements & Influence of the Program

Script: I would like to begin by asking you some questions about the scientific advancements made through the Program over the past five years, and the influence of the Program in the fields of proteomics and heart, lung, and blood research.

5. When you consider the work of the Program, what, if any, significant discoveries or advancements have emerged from the Program that you would consider to be significant in the fields or proteomics and/or heart, lung, and blood research?
    a. How, if at all, would you say these discoveries have influenced proteomics and heart, lung, and blood research?
6. Thinking about the connection between proteomes and the molecular phenotypes of disease, what do we know that we did not know 5 years ago, based on the work of the Program?
    a. In what ways, if at all, has the Program contributed to our understanding of the underlying mechanisms of heart, lung, and blood disease biology?
7. How would you characterize the capacity of the field for innovative research as compared to five years ago?
    a. In your opinion, how, if at all, has the Program increased the capacity of the field as a whole to engage in innovative HLB research?
8. In what ways, if at all, has the Program raised the profile of proteomics research?
    a. Relative to other work being done in the field, how influential has the work of this Program been, in your opinion? For example, would you say that it is on par with the field, leading the field, or lagging behind the rest of the field?

Script: The next questions ask specifically about clinical and translational advancements. I'll start by asking specifically about the work of the Program, then move on to think of the field as a whole.

9. How, if at all, has the Program contributed to the development of knowledge or tools relevant to clinical or translational questions in the field?
10. In your opinion, how has the Program advanced the field in terms of clinical outcomes?
11. Now thinking about the field as a whole, how would you describe the next steps needed in terms of clinical applications of proteomic research? In other words, where are we now, as this five-year phase of the Program is wrapping up?

Section B: Application & Dissemination of New Knowledge and Tools

Script: Great. Now I am going to ask you about the application and dissemination of knowledge and tools developed through the Program.

12. In your work, how, if at all, have you used new knowledge or tools that were developed through the Program?
    a. How or through what channels did you access new knowledge and tools developed through the Program?
13. In what ways, if at all, have you seen knowledge, methods, or tools that emerged from the Program applied by other researchers, or applied elsewhere in the field?
14. In what ways, if any, had you expected or anticipated tools developed by the Program to be made available to the broader research community, that were not realized?

Section C: Collaboration

Script: My next questions focus on the role and nature of collaborations associated with the Program.

15. In what ways, if any, have you or your colleagues collaborated with researchers involved in the Program?

a. How would you characterize these collaborations? For example, with whom did you tend to collaborate, what was the nature of the collaborations, and how would you describe the quality of the collaborations?
b. Of the collaborations in which you were involved, how many represented new collaborative relationships? Previously activated collaborative relationships?

Script: Now I would like to ask a question about the more formal collaborative relationships that may have developed between the Program and other researchers in the field.

16. In what ways, if any, have you or your colleagues engaged in formal collaborations that involved working collectively with researchers from the Program to achieve a common goal?
    a. How would you describe the nature of those collaborations?
    b. In your opinion, how valuable were these collaborations for your work? For example, would you say that they were instrumental to your work, generally productive, or generally unproductive?
17. In what ways, if any, do you intend to leverage or build on these collaborations going forward?

Section D: Recognition & Advancement

Script: My final questions focus on the recognition and professional advancement of researchers involved in the Program.

18. I am going to ask you to consider the work and standing of the lead investigators involved in the Program. Relative to other leading investigators in the field, how would you characterize the leading investigators who are involved with the Program?
    a. For example, would you say they are on par with other investigators, exceptional in the field, or lagging behind other senior investigators in the field?
19. In the past five years, what, if any, major conference or keynote presentations that were based on the work of the Program stand out as being particularly significant in your mind?
    a. Please describe the context and content of the presentations (who, where, what was it about, why it was significant).
20. That is the last of my formal questions. Is there anything about the Centers or the Program that you feel is important for us to know, that I didn't ask about?

Closing Remarks

Script: Before we close, I would like to ask you about your level of familiarity with the Program.

21. Would you say that you have: a) considerable familiarity, b) some familiarity, or c) little or no familiarity with the Program? [Circle one]
    a. How have you learned about the Program? Was it primarily though: a) first-hand involvement, such as relationships, collaborations, or direct professional involvement with Program-affiliated researchers, or b) second-hand involvement, such as publications and conference presentations. [Circle one]
    b. Within the Program, are there particular Centers or researchers with which, or with whom, you are familiar? Please be specific.

Script: Thank you again for taking time out of your schedule to participate in this study. We will be using the data we collect through these interviews to inform our evaluation of the Program. If, after revisiting our interview, I have questions for clarification, is it okay if I contact you again? [YES/NO]
Great. Thank you again for your time. Please feel free to contact us if you have any follow-up questions or comments.

Interviewer notes

Methodological comments
• How did the process go? What worked well? What didn't? • Observations on the questions/guide: (Redundancy; Flow; Specificity/generality of the questions) • Other
Analytical comments
• Thoughts/observations on the content of the interview • Themes or connection • Demeanor of interviewee • Key new information • Other

NHLBI Proteomic Centers Program Study
Key Informant Interview Guide – NIH/NCI/NHLBI

Interview Information

Interviewee:		
Title:		
Interview Format	□ In-person □ Phone □ Skype/video chat	
Phone/Email:		
Video chat ID:	Skype:	GooglePlus:
Date:	Start time:	End time:
Interviewer:		

Introduction

Opening Script:

- Thank you for taking the time to participate in this study.
- My name is _____, and I am with Concept Systems, Inc. We are an evaluation firm that has been contracted to conduct an evaluation of the NHLBI Proteomics Centers Program (which I will refer to as "the Program" from now on).
- Goals of the evaluation are to better understand the contributions of the Program, and the experiences of Program staff and affiliates.
- Looking at the Program as a whole. Findings will not be used to compare the individual Centers.
- Interview should take about 45 minutes.
- This evaluation has been approved by the OMB. Not research, so IRB approval not required. However, your participation is voluntary – can skip questions or stop.
- The evaluation only covers the past five years of the Program, from 2010 to 2015. If you were involved in previous iterations of this program, please make sure that your responses reflect your experience only over the past five years.
- Info you provide will not be disclosed to anyone but the researchers conducting the study, except otherwise required by law.

With that said, are you willing to continue with the interview? [YES/NO].
And is it okay if I audio record the interview? [YES/NO]
Do you have any questions before we begin? Let's get started.

Background Questions

Script: First, I want to gather some basic information about your role and position.

1. Please verify the following background information:

Title:	
Center/Organization:	

2. In what capacity have you worked with the Program?
3. In responding to my questions, please remember that we are focused only on the past 5 years of the Program. Has your role with the Program changed during that time?
4. Very briefly, how would you describe your current role?

Script: Before we begin, I would like to ask you about your level of familiarity with the Program.

5. Would you say that you have: a) considerable familiarity, b) some familiarity, or c) little or no familiarity with the Program? [Circle one]

a. How have you learned about the Program? Was it primarily though: a) first-hand involvement, such as relationships, collaborations, or direct professional involvement with Program-affiliated researchers, or b) second-hand involvement, such as publications and conference presentations. [Circle one]
b. Within the Program, are there particular Centers or researchers with which, or with whom, you are familiar? Please be specific.

[Note: If only familiar with Centers, rather than Program, adjust questions accordingly. Advise respondent to give response about the Program, to the best of their knowledge.]

Section A: Scientific Advancements & Influence of the Program

Script: I would like to begin by asking you some questions about the scientific advancements made through the Program over the past five years, and the influence of the Program in the fields of proteomics and heart, lung, and blood research.

6. When you consider the work of the Program, what, if any, discoveries or advancements have emerged from the Program that you would consider to be significant in the fields of proteomics and/or heart, lung, and blood research?
    a. How, if at all, would you say these discoveries have influenced proteomics and heart, lung, and blood research?
7. Thinking about the connection between proteomes and the molecular phenotypes of disease, what do we know that we did not know 5 years ago, based on the work of the Program?
    a. In what ways, if at all, has the Program contributed to our understanding of the underlying mechanisms of heart, lung, and blood disease biology?
8. How, if at all, has the Program increased the capacity of the field as a whole to engage in innovative proteomics and heart, lung, and blood research?
    a. What role, if any, has collaboration played in influencing the capacity of the field?
9. In what ways, if at all, has the Program raised the profile of proteomics research?
    a. Relative to other work being done in the field, how influential has the work of this Program been, in your opinion? For example, would you say that it is on par with the field, leading the field, or lagging behind the rest of the field?

Script: The next questions ask specifically about clinical and translational advancements. I'll start by asking specifically about the work of the Program, then move on to think of the field as a whole.

10. How, if at all, has the Program contributed to the development of knowledge or tools relevant to clinical or translational questions in the field?
11. In your opinion, how has the Program advanced the field in terms of clinical outcomes?
12. Now thinking about the field as a whole, how would you describe the next steps needed in terms of clinical applications of proteomic research? In other words, where are we now, as this five-year phase of the Program is wrapping up?

Section B: Application & Dissemination of New Knowledge and Tools

Script: Great. Now I am going to ask you about the application and dissemination of knowledge and tools developed through the Program.

13. In your work, how, if at all, have you used new knowledge or tools that were developed through the Program?
    a. How or through what channels did you access new knowledge and tools developed through the Program?
14. In what ways, if at all, have you seen knowledge, methods, or tools that emerged from the Program applied by other researchers, or applied elsewhere in the field?

15. In what ways, if any, had you expected or anticipated tools developed by the Program to be made available to the broader research community, that were not realized?

Section C: Collaboration

Script: My next questions focus on the role and nature of collaborations associated with the Program.

16. In what ways, if any, have you or your colleagues collaborated with researchers involved in the Program?
    a. How would you characterize these collaborations? For example, with whom did you tend to collaborate, what was the nature of the collaborations, and how would you describe the quality of the collaborations?
    b. Of the collaborations in which you were involved, how many represented new collaborative relationships? Previously activated collaborative relationships?

Script: Now I would like to ask a question about the more formal collaborative relationships that may have developed between the Program and other researchers in the field.

17. In what ways, if any, have you or your colleagues engaged in formal collaborations that involved working collectively with researchers from the Program to achieve a common goal?
    a. How would you describe the nature of those collaborations?
    b. In your opinion, how valuable were these collaborations for your work? For example, would you say that they were instrumental to your work, generally productive, or generally unproductive?
18. In what ways, if any, do you intend to leverage or build on these collaborations going forward?

Section D: Recognition & Advancement

Script: My final questions focus on the recognition and professional advancement of researchers involved in the Program.

19. I am going to ask you to consider the work and standing of the lead investigators involved in the Program. Relative to other leading investigators in the field, how would you characterize the leading investigators who are involved with the Program?
    a. For example, would you say they are on par with other investigators, exceptional in the field, or lagging behind other senior investigators in the field?
20. In the past five years, what, if any, major conference or keynote presentations that were based on the work of the Program stand out as being particularly significant in your mind?
    a. Please describe the context and content of the presentations (who, where, what was it about, why it was significant).
21. That is the last of my formal questions. Is there anything about the Centers or the Program that you feel is important for us to know, that I didn't ask about?

Closing Remarks

[Revisit familiarity question, Q5, if needed]

Script: Thank you again for taking time out of your schedule to participate in this study. We will be using the data we collect through these interviews to inform our evaluation of the Program. If, after revisiting our interview, I have questions for clarification, is it okay if I contact you again? [YES/NO] Great. Thank you again for your time. Please feel free to contact us if you have any follow-up questions or comments.

Interviewer notes

Methodological comments
• How did the process go? What worked well? What didn't? • Observations on the questions/guide: (Redundancy; Flow; Specificity/generality of the questions) • Other
Analytical comments
• Thoughts/observations on the content of the interview • Themes or connection • Demeanor of interviewee • Key new information • Other

C.9. Quality Control Checklist

Quality Control Checklist **NHLBI Key Informant Interviews** Version Date: 01.16.15		
The purpose of this checklist is to provide a set of specific elements necessary for a consistent and thorough interview process. The checklist is meant to serve as a quality review tool to be completed by an external reviewer during the interview process.		
1.0	**OPENING & INTRODUCTION**	
☐	1.1	Interviewer has introduced self
☐	1.2	The purpose of the interview has been explained
☐	1.3	The statement of recording has been reviewed
☐	1.4	The statement of confidentiality and informed consent has been reviewed
☐	1.5	The respondent's consent to proceed has been obtained.
☐	1.6	Interview information section of the interview guide has been completed.
2.0	**INTERVIEW CONDUCT**	
☐	2.1	Respondent's background has been verified.
☐	2.2	Interviewer uses probes and prompts, where appropriate.
☐	2.3	Interviewer follows up/seeks clarification, where appropriate.
☐	2.4	Interviewer engaged in active listening.
☐	2.5	Interviewer sought clarification for acronyms or unfamiliar terminology.
☐	2.6	Interviewer managed the pace and timing of questions appropriately.
☐	2.7	Interviewer matches questions with discussion content and interview flow.
3.0	**CLOSING**	
☐	3.1	Expression of gratitude for respondent's time has been offered.
☐	3.2	The next steps of the inquiry have been relayed.
☐	3.3	Follow-up with the respondent has been addressed.
4.0	**DOCUMENTATION**	
☐	4.1	Interviewer made handwritten notes during the interview
☐	4.2	Interviewer completed methodological and analytical sections at the end of the interview guide
☐	4.3	Interviewer completed post-interview summary/memo
5.0	**TECHNOLGY**	
☐	5.1	Two digital recording copies have been made, appropriately named and saved to the O drive
☐	5.2	Interview platform (Skype, phone) functioned properly (if applicable)

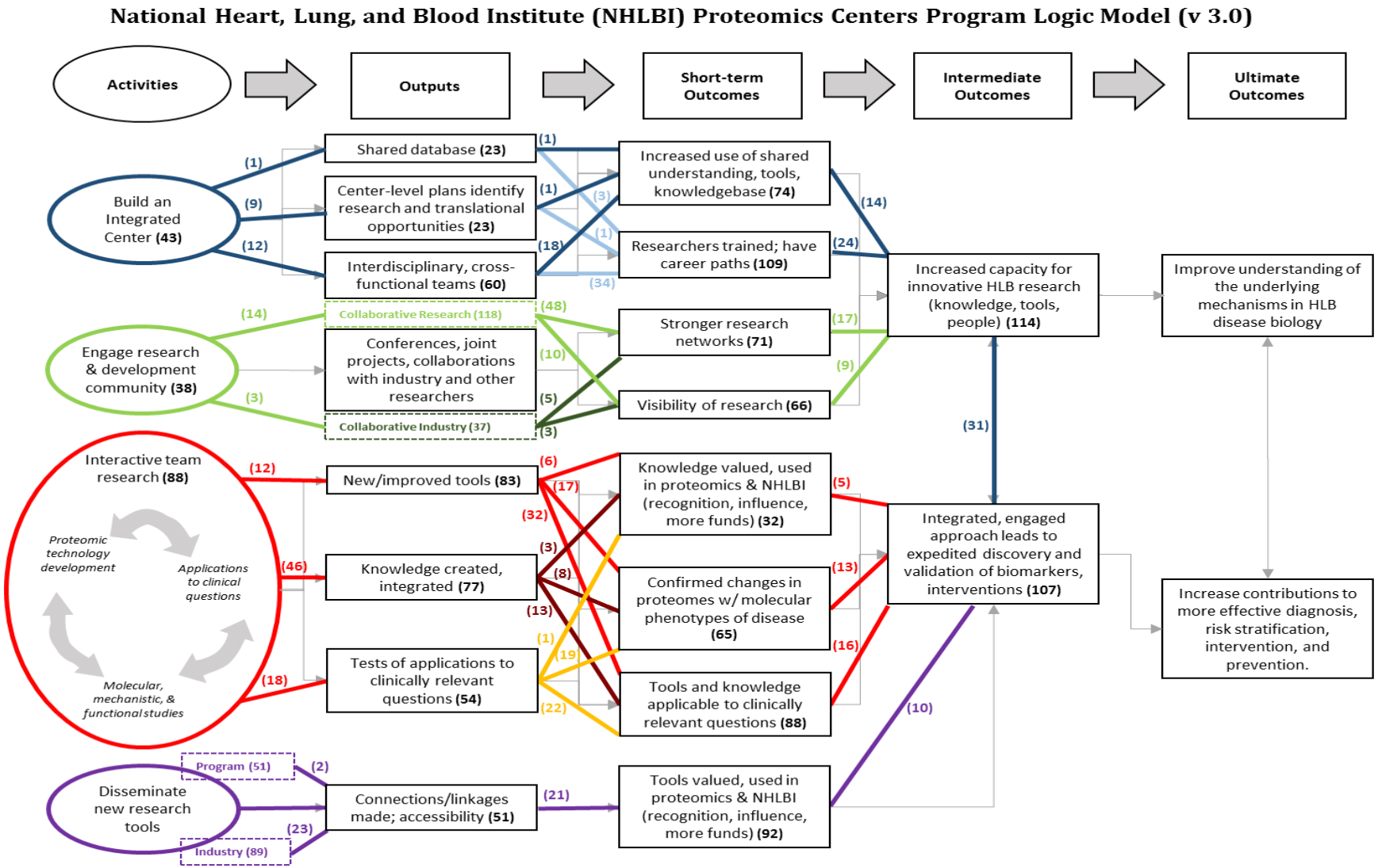
C.10. Analysis Codebook

1. Shared database	**SHARED_DATABASE**	**Creation and use of a shared database within the Center**
Node	Code	Description
2. Build an integrated center	INTEGRATED_CENTER	Formation and bringing together of people and resources for the center.
3. Center-level plans identify research and translational opportunities	PLANS_FOR_RESEARCH	Stated plans or goals for research within the Center (purpose of the Center)
4. Interdisciplinary, cross-functional teams	INTERDISCIPLINARY_TEAMS	Data related to interdisciplinary teams being established or coming together at the Centers (*separate* from active interdisciplinary research activities)
5. Increased use of shared understanding, tools, knowledgebase	USE_OF_SHARED_TOOLS	Center or Program level examples of how knowledge, tools are shared and applied within the Program
6. Interactive team research and training	INTERACTIVE_TEAM	Role, function, and mechanics of the interactive team approach. Active team integration, research activities, etc.
7. Researchers trained	RESEARCHERS_TRAINED	Researchers trained and have career paths.
a. Researchers trained: have career paths	ADVANCEMENT	Evidence of career advancement, changes in or reflections on career path
8. New/Improved tools	NEW_TOOLS	Tools created or improved
9. Knowledge created, integrated	KNOWLEDGE_CREATED_INT	New information being generated AND integrated across disciplines and areas. *How learning is being manifest in this integrated model*. (Incl. new discoveries that open up the door to inquiry, such as someone who didn't see the value of proteomics, and is now writing it into his/her grants).
10. Tests of applications to CLINICAL questions	TEST_APPL_TO_CLINICAL_QUESTIONS	Generation of knowledge that can be applied clinically. Mechanics. Dynamic learning and applying in clinical contexts. E.g., heart attack data sent to PIs within the Program.
11. Tools valued	TOOLS_VALUED	Tools valued or used by others in the fields of proteomics/HLB research, *incl. additional funding leveraged AND influence on the field*

12. Knowledge valued	KNOWLEDGE_VALUED	Knowledge valued by others in the field, OR evidence of peer acknowledgment [FOR INDIVIDUALS], including awards and other forms of recognition
13. Discoveries: Confirmed changes in proteomes with molecular phenotypes of disease	LINK_PROTEOMES_W_DISEASE	Linkages confirmed between proteomes and specific phenotypes of disease
a. New or revised scientific aims	NEW_REVISED_SCIENTIFIC_AIMS	New or revised scientific aims within the Center
14. Discoveries: Tools and knowledge applicable to clinical questions	CLINICAL_RELEVANCE	References to tools and knowledge applicable to clinical questions
a. Clinical goals	CLINICAL_GOALS	Stated clinical goals of the Center (or lack of knowledge of clinical goals)
b. Clinical advancements from the PROGRAM	CLINICAL_ADVANCEMENTS_PROGRAM	Responses to, what clinical advancements have emerged from the Program as a whole?
c. Clinical next steps for the FIELD	CLINICAL_NEXT_STEPS_FIELD	Next steps for the field in terms of clinical work/applications.
15. Disseminate new research tools/prototypes [within the Program]	DISSEMINATE_PROGRAM	Dissemination of knowledge, tools, etc. WITHIN the Program (either direction)
16. Disseminate new research tools/prototypes [outside the Program]	DISSEMINATE_FIELD	Dissemination of knowledge, tools, etc. OUTSIDE the Program (from Center/Program to field), *including publications*.
17. Connections/linkages made; accessibility	CONNECTIONS_LINKAGES	How to bring together researchers, technologists, industry, to really drive the development of the science. Didn't exist before, so had to build the plane while flying it. That required *bringing everything together. Broad outreach, and tying everything together in extensive networks*. (Exploratory; Bigger than collaborations)
18. Engage research and development community	ENGAGE_R-D_COMMUNITY	How is information from NHLBI getting *out into the world*? *Who is accessing the knowledge and data*?
19. [Value of] Joint projects, collaborations with *other researchers*	COLLAB_RESEARCH	Perceived value of joint projects/collaborations with other researchers (Program or field)
20. [Value of] Collaborations with *industry*	COLLAB_INDUSTRY	Perceived value of collaborations with industry

21. Stronger research networks	RESEARCH_NETWORKS	Outgrowth, extension, leverage of collaborations (next steps of collaboration)
22. Visibility of research	VISIBILITY	Anything that makes the work of the program visible or accessible to the world (podcasts, etc.). OR lack of knowledge of the program by outsiders
23. Increased capacity of innovative proteomics/HLB research	CAPACITY_OF_FIELD	Responses to the capacity question, or other references to field expansion
24. Expedited discovery and validation of biomarkers, interventions	EXPEDITED_DISCOVERY	Examples of how the Program has expedited discovery within the field. Also, examples of how the field has rapidly progressed in the past five years.
EMERGENT		
a. (25) Funding mechanism	MECHANISM	References to the contract/funding mechanism
b. (26) Uniqueness of Center within Program	CENTER_UNIQUE	Perceptions that one's center is unique from all other centers (e.g., we can't really collaborate with the other centers, because what we do is so different)
c. (27) State and role of technology	STATE_AND_ROLE_OF_TECHNOLOGY	What is the current state of technology in the field? Is it sufficient for the work that needs to be done, or insufficient? Where is it lagging/exceeding the other areas?
d. (28) Democratization of the science	DEMOCRATIZATION_OF_THE_SCIENCE	Knowledge and tools should be made broadly accessible to everyone
e. (29) NHLBI Leadership	NHLBI_LEADERSHIP	Comments on leadership of the Program or NHLBI/NIH
f. (30) Gender	GENDER	Comments on gender, including the role of women in leadership in the Program
g. (31) Narrow focus	NARROW_FOCUS	Comments on the Program being too narrow in focus
GUIDE		
a. (32) Constraints to work	CONTRAINTS	Things that constrain the work of the Centers or Program
b. (33) Global assessment of Program's lead investigators	ASSESS_INVESTIGATORS	Responses to: How would you characterize the lead investigators within the Program?
c. (34) Global assessment of the Program as a whole	ASSESS_PROGRAM	Responses that reflect general assessments of the value or quality of the program (generally by outsiders).
d. (35) Influence from the field on the Program/Centers	INFLUENCE_FROM_FIELD	Examples of how the centers or program have been influenced by research/advances elsewhere in the field

C.11. Codebook Matrix



Appendix D. Data Abstraction/Scientific Literature Review, Bibliometrics, Altmetrics, and Publication Analysis

D.0. Scientific Publication Data Protocols and Procedures

The evaluative of scientific publications seeks to assess the impact of scientific output in the context of other published science and usually compares the relative scientific contributions of research groups or institutions. The publication analysis involves accessing, organizing, and analyzing scientific publications and product output. This type of analysis requires the initial development and implementation of a structured process to manage and verify the source material. From an accurate source of research publication output, two primary methods for assessing research recognition were used to obtain data in relation to indicators of productivity and linkage: bibliometrics and altmetrics. A standardized protocol was developed to extract publication data at the paper-level, journal-level and field-level. In addition to the bibliometric and altmetrics publication analysis, a comprehensive coding analysis was conducted through web analytics, annual reports, and the biosketch and/or curriculum vitae of key personnel in the Program. This allowed for a robust and thorough examination into all areas of scientific contributions from those in the Centers.

D.1. Bibliometric Protocol

The primary assumption that supports the use of bibliometrics is that exchange and recognition of research results is desired and is one of the key driving forces in the advancement of science. Citations symbolize the association of scientific ideas, and the references which authors cite in their papers make explicit the link between their current research and prior work in the scientific literature archive.[19] While not a definitive statement of impact, these data are supplemented with inferential analyses to yield a picture of how the center publications are influencing subsequent peer-reviewed science and other research across proteomic relevant fields.

The publication analyses included examination of citation patterns, network relationships, and online uptake using altmetrics. The analytic approach involved descriptive, inferential and analysis of patterns including the following: recognition, sharing/linkage, subject/category contributions, prominence/visibility, collaboration, and cross-method linkage. The information generated through these analyses is intended to provide insight into how the Center papers published over the past five years are performing relative to the larger body of scientific literature.

D.1.1. Publication Abstraction

In order to prepare the publications to collect (extract) the data, an initial list of publications from each of the Centers were provided to the evaluation team through the NHLBI COR. The lists consisted of compiled submitted publications within the specified time period between August 2010 and August 2015. This produced 1,040 papers published. These lists were reviewed and organized. Then each bibliographic record on the paper was verified in PubMed. This process led to the creation of a 706-item PubMed collection where the complete list of verified bibliographic records of the Centers research output were maintained.

Utilizing the verification process described, these papers in PubMed yielded 1,040 unique PubMed Identification Number (PMID) numbers. A batch search of the 1,040 unique PMID numbers in the Web of Science (WoS) yielded 1,008 matches and an individual search matched the remaining 32 papers. At the paper-, journal -, and field-level, metrics for the final list of 1,040 papers was extracted from Incites software[20], where 975 (93.8%) of the Centers papers were found to have accumulated citation data.

[19] Koskinen J, Isohanni M, Paajala H, et al. (2008). How to use bibliometric methods in evaluation of scientific research? An example from Finnish schizophrenia research. *Nordic Journal of Psychiatry.* 62:136–43.

[20] Incites website (accessed 1 Jun 2015): http://thomsonreuters.com/en.html

In addition to the initial list of publication provided, 342 publications and records were extracted by the final annual reports from each of the seven Centers. As part of the standardized protocol, publications were initially searched using PubMed and Google Scholar. Papers with identified PMID numbers were then verified in the Web of Science (WoS). Papers without PMIDs were manually searched in the WoS and publication data was obtained if located.

To ensure seamless data management and gathering, Eigenfactor.org, Thomson Reuters, and Incites Research Analytic Platform were used. In particular, journals in which the Center papers have been published were identified using Eigenfactor, which allowed collection of metrics related to the impact and prominence of the journals (i.e. Integrated Impact Factors and Eigenfactors). Furthermore, Eigenfactor was utilized to establish bibliometric indicators by fields (i.e. ISI categories). Incites was specifically used to collect Essential Science Indicators related to paper-, journal-, and field- level metrics.

D.1.2. Bibliometric Analyses Measures

Analyses were conducted for the Centers publication record for the period 2010-2015. Analyses included assessment in five primary areas:

a) performance (the initial output metrics based on citation counts);
b) peer-review presence (the status of a set of papers based on the ranking of journals in certain fields);
c) relative impact (the normalized comparisons of a set of papers to others published in similar journals and fields);
d) collaboration in the form of co-authorship patterns over time; and
e) online presence through uptake and distribution in social media.

The analysis covers 975 Centers' papers, published by 4,379 unique authors, in 285 journal, across 70 fields, and spans collaborations in 49 countries, including the United States. Several metrics for publications by article, research articles and other papers (e.g. reviews, editorials, letters, proceedings papers & meeting abstracts) were analyzed using the following metrics:

1) Total papers
2) Total citations
3) Mean times cited
4) Percent of papers cited
5) C-index
6) H-index

The C-index describes how the group of papers preformed overall and equals the sum of all actual citations divided by the sum of all the expected citations. This calculation indicates a ratio of actual to the expected number of citations for a group of papers, with an index of 1.0 indicating that all papers in the category received the average number of citations for their respective journal, article type, and year. The H-index is a statistic used to measure both the scientific productivity and the apparent impact of a selected category and is measured by the number of papers (N) that has N or more citations.

D.1.3. Yearly Performance and Impact

The evaluation team examined several metrics of performance and impact overall, as well as by each year to gauge how well the Centers' publications were performing relative to journals and fields where they have been published. The analyst followed the updated procedures and terminology outlined by Center

for Science and Technology Studies at Leiden University for assessing impact of a set of papers relative to those published worldwide.[21,22]

First, the analyst calculated the total number of citations for the set (Total Citation Score; TCS) and citations per paper (Mean Citation Score; MCS) by year and publication type. Next, the analyst assessed:

a) the average total number of citations of a certain article type (abstract, article, review, note, etc.) published in a specific journal cumulatively by the most recent completed year and;
b) the average number of citations for all articles that were published in a particular year, in all journals in a specific field from the Journal Citation Reports (JCR) database. The analyst then summed all of the corresponding journal specific values for each paper and calculated a journal normalized measured impact ratio (Mean Normalized Journal Score; MNJS).

Finally, the analyst summed all of the corresponding field-specific values for each paper and calculated a field normalized measured impact ratio (Mean Normalized Citation Score; MNCS). In both cases, normalization of the citation values is completed at the paper level that corresponds to the selected publications with respect to the publication type, age, and journal or subject area. The Integrated Impact Factor (I3) was also calculated.[23] The I3 applies non-parametric statistics using percentiles, allowing highly cited papers to be weighted more than less-cited ones. The indicator is built upon the premise that impacts add up instead of average out. The evaluation team further evaluated the global standing of the journal set containing the papers from the Centers relative to the fields in which the journal belongs.

D.1.4. Paper-Level Analyses

Descriptive analyses were preformed to assess network publications based on the number of highly cited papers at various percentiles. This measure is a breakdown of number of papers that have the necessary number of citations to meet highly cited thresholds of the top 1%, 5%, 25% and 50% of publications in their respective fields.

D.1.5. Journal-Level Analyses

Journals are not homogeneous outlets of science in terms of their audiences, visibility, significance, and readership. Across fields of research, great value is placed on journals with higher status and perceived levels of productivity and therefore attracts large international audiences from the scientific community. Journals with the propensity to draw a great deal of attention to the papers it publishes are held in high regard and widely recognized across international settings. Requirements for publishing in these journals are often stringent and review processes are particularly demanding. Therefore, analyses that reveal where a set of papers reside in the journal hierarchy is important to evaluating the presence of a set of papers across scientific fields. The JCR Impact Factor is a well-known metric in citation analysis. It is a measure of the frequency with which the "average article" in a journal has been cited in a particular year. The impact factor helps evaluate a journal's relative importance, especially when compared to others in the same field. The impact factor is calculated by dividing the number of citations in the JCR year by the total number of articles published in the two previous years. An impact factor of 1.0 means that, on average, the articles published one or two years ago have been cited one time. An impact factor of 2.5 means that, on average, the articles published one or two years ago have been cited two and a half times. Citing articles may be from the same journal; however, most citing articles are from different journals.

[21] Waltman L, Van Eck NJ, van Leeuwen TN, et al. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics.* 5(1):37-47.
[22] Waltman L, Van Eck NJ, van Leeuwen TN, et al. (2011). Towards a new crown indicator: An empirical analysis. *Scientometrics.* 87(3):467-481.
[23] Wagner CS, Leydesdorff L (2012). An integrated impact indicator: a new definition of 'impact' with policy relevance. *Research Evaluation*. 21(3):183-188.

Journals were examined in which network papers were published. The 975 papers were published in 285 different journals, with the number of publications ranging from 45 to 1. Using the JCR Impact Factor as means for organizing the journal hierarchy and evaluating the position and presence of network publications, the top tier of journals was examined for which network papers were found.

D.1.6. Field-Level Analyses

In an effort to examine the performance and presence of Centers' papers in relation to a wide variety of scientific areas of study, the evaluation team conducted an analysis of bibliometric indicators by fields (i.e., ISI Categories). Specifically, within each field, the evaluation team examined the volume of network papers linked to each field, the level at which papers were published in each field according to the journal rankings in each field, and the proportion of the field specific set at different ranking thresholds.

The full set of 975 Center papers were distributed across 70 specific fields. As a means for understanding the level of the network papers relative to the importance of the journals within each field, we ranked ordered the journals in each field by Eigenfactor score and identified the top 25% (1st quartile) as an indication of the top tier of journals for that respective field. The Eigenfactor score is a measure of the total influence of a journal based on cross-citation patterns. The Eigenfactor uses the 5 previous years for the target window and excluded journal self-citations.[24] Thus, the top 25% of journals in a field, based on Eigenfactor ranking, is a robust indicator of the prominence of the journal within the field. The number of papers that were published in the top 3 ranked journals were also counted in their respective field, further distinguishing the highest level of recognition within each field.

D.1.7. Profile of Interdisciplinary Research Analysis

An important part of advanced bibliometric performance and impact evaluation is the construction and representation of a research profile for a specific entity.[25,26] A research profile is a breakdown of output, performance, and impact according to internationally defined research fields on the basis of the journals used by the entity. The profile of interdisciplinary research, and computed the normalized metrics for worldwide comparison, with focus on only those papers considered "research articles" were analyzed. In order to fully understand the relative differences within and across fields, the normalized (i.e., adjusted for age, type, and source) metrics of worldwide impact were analyzed.

D.1.8. Co-authorship and Collaboration Analyses

Network analyses were conducted to visually represent and assess the collaborative connections of the research output from the Centers using VOSviewer.[27] The software was used to conduct the visualization analyses and Gephi[28] to compute the separate network metrics. The initial step in computing the co-author networks was to disambiguate author names resulting from misspellings, use of different initials, and hyphenated names. A thesaurus file was developed and used in the analysis in order to combine different author versions and correct misspellings.

In addition, Center personnel, external experts connected to the program, and NIH personnel involved in the management and evaluation were identified with a unique character attached to their name. Beginning with 2010, the author lists from each subsequent year was added to the previous set to account for the growth and expansion of the co-author relationships. The shared relationships were determined by counting the instances where individuals listed as co-authors on the same publication. For each yearly slice of co-authorship data, the visualization parameters in VOSviewer (i.e. mapping attraction and

[24] Bergstrom CT, West JD, Wiseman MA (2008). The eigenfactor metrics. *Journal of Neuroscience*. 28:11433-34.
[25] Van Raan, AFJ (2008). R&D evaluation at the beginning of a new century. *Research Evaluation*. 8:81–6.
[26] Van Raan AFJ (2005). Measurement of central aspect of scientific research: performance, interdisciplinarity, structure. *Measurement.* 3:1–19.
[27] Van Eck NJ, Waltman L (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics.* 84(2): 523-538.
[28] Bastian M, Heymann S, Jacomy M (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.

repulsion, normalization method, cluster resolution) were consistent across representations so as to generate comparable network visualizations. In addition, for each year the normalized adjacency matrix for each network map visualized in VOSviewer was analyzed in Gephi to generate common network metrics. These network metrics included:

1. Nodes (Number people in the measured network)
2. Edges (Number of links or ties in the measured network)
3. Average Degree (Average number of ties between network members)
4. Average Weighted Degree (Average strength of the ties between network members)
5. Density (Measure of how well connected the network is or how "close knit"; higher value = more densely connected)
6. Modularity (Measure of the density of connections between members within and across modules (i.e., groups or communities)
7. Average path length (Average number of steps it takes to get from one member of the network to another)

D.1.9. Scientific Aims

The scientific publications output in relation to the scientific aims established by the Centers was examined as a proxy for the program research agenda. The scientific aims for each Center were provided by the NHLBI Project CORs. An initial coding of each Center's scientific aims by specific focus, subject, and study was performed. This coding enabled efficient linkage of research output produced by the Center with appropriate scientific aim.

Using the publication set, 62 scientific categories were identified in the WoS describing the breadth of the scientific content areas of the Center's output. These standardized categories (assigned by ISI) enabled further organization of scientific content and were linked to scientific aims for a precise analysis of the research output. Furthermore, using the author supplied keywords found on the publication submitted by the centers, we analyzed the co-occurrence of key words as descriptors of the content of the articles. The analysis produced an adjacency matrix on the frequency of keyword co-occurrence. Using this adjacency matrix, we computed a density visualization in VOSviewer that partitioned the keyword content into a network of keywords, in effect representing a general content analysis of the scientific publication corpus supplied by the center. This representation was compared against the scientific agenda of the program, as represented by the scientific aims of the centers.

D.2. Altmetric Protocol and Analysis

Altmetrics comprise metrics based on the integration of social media tools that can inform broader, faster measures of impact, and can serve to complement traditional citation metrics.[29] In this regard, altmetrics were used to describe how the Centers' research output is being publicly shared via social media platforms. In accessing altmetric data for the Centers publications, Altmetric Explorer was used. This produced approximately 12,000 online mentions of individual scholarly articles are found every day.[30] This web-based platform includes comprehensive access to social media (i.e. Pinterest, Facebook, Twitter, Google+), reference managers (i.e. Mendeley, CiteULike), blogs, and mainstream media outlets.

To incorporate altmetrics, the evaluation team identified several "event categories" that are appropriate for the research outputs of the Centers, with events defined as a specific actions applied to articles, such as bookmarks or tweets. Using the PMID numbers of the papers published by the Centers, the evaluation team queried the Altmetric Explorer database for event counts across several categories. Event counts were simply the total number one certain type of event to date. Altmetric Explorer provides percentile

[29] Liu J, Adie E (2013). New perspectives on article-level metrics: developing ways to assess research uptake and impact online. *Insights.* 27(2): 153-158.

[30] Altmetric Explorer website (accessed 1 Mar 2016): https://www.altmetric.com/

scores for all articles, and adjusted for journal and year of publication. Summaries for the event counts by categories were produced and represented.

D.3. Web Analytics

Web-analytics were collected to inform dissemination and uptake patterns associated with the web-based platforms where the program's scientific research output and information was housed. Specifically, these data were obtained by the Program's Administrative Coordinating Center to help establish the Program's push and pull dissemination strategies as well as the following:

1. the visibility of the research and its presence in the media;
2. how Center teams are accessing new knowledge/tools;
3. how the tools/technologies developed by the Centers are accessed by other researchers; and
4. how those tools/technologies are made available.

The following data were collected by the Administrative Coordinating Center:

1. Web analytic data for the main program website and salient components of the website (pull strategy), by:
    a. Unique visitors
    b. Page views
    c. Average page views by visit
    d. Countries/cities (origin of access)
2. Summaries of the distribution of program products (push strategy).
3. Listing of the conference locations (push strategy).
4. Connections with the scientific organizations and communities (push strategy).

Summaries based on the data received from the Administrative Center at UCLA were generated by year and represented.

D.4. Annual Reports Procedures

Between February 2015 and September 2015, qualitative and quantitative data were collected and extracted from the final annual progress reports of each Center. This allowed evaluators to describe both the progress made to date and plans for the following year. A collection of the Centers semi-annual and annual reports during the period of 2010 – 2015 were provided by the NHLBI CORs and program staff.

An inventory database of the report collection was developed and receipt of the reports were tracked accordingly. In preparation for analysis, a structured protocol for review and coding was developed using OSR NVivo Data Analysis Software (NVivo) to facilitate extraction of key data that is specific to Centers as well as general across the enterprise. The evaluation team extracted and coded descriptions of the accomplishments of the research project, including aims, studies and results, significance, plans, and publications and other project generated resources. The evaluation team was also able to extract 342 publications and records, which were included in the final publication sample.

The analysis of the annual reports included content analysis with application of the codebook developed for the analysis of the qualitative interviews **(See Appendix C: Key Informant Interviews and Qualitative Analysis)**. In particular, qualitative interview codes consistent with the indicators operationalized in the logic model were used to inform the coding process. This coding process allowed purposeful examination of qualitative material related to the Centers' clinical goals and plans for research.

In order to maximize reliability and ensure that data elements were populated consistently with expectations outlined in the protocol, strict quantitative quality control and data audit procedures were implemented. To ensure that quality assurance was applied to each document, all quality checks were

tracked and monitored using the developed tracking database. The goals of the quality control audit were to ensure that:

- All protocols were abstracted completely
- Determine potential missing data trends
- Ensure accurate data collection

As part of the quality assurance process, a standardized memo approach and missing data audit were implemented. After each abstraction, the abstractor completed a memo noting missing data, analytical and methodological reflections. Subsequent memos were reviewed by the data abstraction task lead and reconciled as necessary with the internal evaluation team. In terms of missing data audits, all completed protocols were reviewed for completeness by the Senior Evaluator. If data was not able to be abstracted for a specific section, the quality reviewer conducted a brief check of the biosketch/CV to confirm this was correct following the checklist found in ***D.6. Biosketch/CV Quality Control Checklist.***

D.5. Biosketches and Curriculum Vitaes Protocol

Between February 2015 and March 2016, the evaluation team collected biosketches and curriculum vitaes (CVs) electronically of key personnel from the Program. A database was developed to manage and organize the collection of biosketches and CVs. This database utilized the person identifier codes consistent with the interview database. This also served as a checks and balance tool in that key personnel identified during the biosketch/CV collection were then cross-checked with the key personnel identified during the interview process. Inconsistencies were reviewed internally, and subsequent requests for biosketches were made or exclusions were determined. The final protocol is found in ***D.7. Biosketch/CV Protocol.***

In preparation for the collection of biosketches and CVs, a multistep approach was developed and implemented to ensure efficiency and accuracy in the collection process. In order to identify key personnel of the program, the evaluation team first conducted an internal search utilizing the Program website and online newsletters as sources. The search revealed 71 total individuals as potential key personnel of the program. To ensure accuracy and completeness of the list of key personnel, three verification steps were implemented:

- First, the list was reviewed by the NHLBI CORS.
- Second, the qualitative interviews database was used to cross check the list of personnel.
- Third, during the collection process, Center representatives were asked to identify additional relevant key personnel. This included 82 key personnel from 3 categories: Directors (7), Co-PI's (19), and Investigators (56).

Due to the availability and content, biosketches were the preferable data source. However, CVs were provided in some instances secondary to availability of material. An inventory database was created to manage and track receipt of collected documents. A total of 14 CVs and 61 biosketches were collected and examined. A biosketch or CV was collected for each of the Directors and Co-PI's, and for 49 of the 56 Investigators **(Table 12 below)**.

Table 12. Biosketch and CV Collected

Center	Documents Received compared to Identified Key Personnel	Biosketches	CVs
Boston	9 of 11	8	1
Harvard	8 of 8	5	3
Johns Hopkins	12 of 13	10	2
Stanford	10 of 10	10	0
UCLA	10 of 13	9	1
UT @ Galveston	14 of 14	13	1
UT @ San Antonio	12 of 13	6	6
TOTALS	**75 of 82**	**61**	**14**

Based on the indicators operationalized in the logic model, a structured protocol was developed to abstract specific data elements from each of the biosketches and CVs in a consistent form to prepare for analysis. The protocol was tested with one biosketch and one CV, which were randomly selected. Findings from the testing phases were reviewed during internal management meetings, and refinements were made to the abstraction protocol as necessary. The finalized protocol was again tested on a random sample of three biosketches and three CVs representing different Centers.

The final protocol was applied to 75 documents, which included the following material:

1. Date of biosketch or CV development
2. The Center
3. Person ID code
4. Title of the file being abstracted
5. Name of the key personnel
6. Year they received their highest degree to determine the length of time in the field of proteomics
7. Information regarding their position advancements
8. Honors and service
9. Non-NIH funding (both ongoing and completed)

Next, a coding system was developed and implemented for each sections of the above data collection area. Identified codes were reviewed and discussed for inconsistencies. The codebook can be found in ***D.8. Biosketch/CV Codebook***.

D.5.2. Biosketches and Curriculum Vitaes (CVs): Advancements, Honors and Services, and Funding

Data related to personal advancements, honors and services, and NIH and non-NIH funding were analyzed to determine the extent to which the program served as a springboard for subsequent academic appointments and professional recognition.

Data were extracted using the following order: advancements, honors and services, and funding information.

- For information regarding their position and advancement, 19 position types were identified and coded. The coding surfaced three categories of positions, which include: Academic, Research and Medical.
- For honors and services, 16 types were identified and coded.
- For non-NIH funding, 7 codes were identified, which included non-NIH organization types. As a final step, if the award amount was available, it was converted to US Dollars using the award date to determine the appropriate conversion rate.[31]

[31] OANDA Currency Converter website: https://www.oanda.com/currency/converter/

The data coding structure described above identified subsequent data that occurred during and prior to the contract period. In particular, the coding process identified advancement position types, categories of positions (i.e. Research, Medical or Academic), types of honors and services, NIH funding award amounts, and associated roles, and non-NIH funding identified by organization type. Data pertaining to advancements were collected for 2010 – 2015. However, honors and professional service, and non-NIH funding data were collected with inclusion of items with designated date ranges that occurred prior to 2010 that extended into 2015, or began in 2015.

NIH-funding data were collected by conducting a search in NIH RePORT. This query included using the first and last names of each identified key personnel with filters indicating funding data available for the period of 2005 – 2015. In cases where the NIH RePORT search data were not found, the biosketches and CVs were examined to determine whether the key personnel reported federal funded studies. After completion of the data abstraction for each biosketch and CV, the data abstractor documented methodological and analytical notes, which were reviewed by the evaluation team, and adjustments were made as necessary.

D.6. Biosketch/CV Quality Control Checklist

Quality Control Checklist **Biosketch and Curriculum Vitae Data Abstraction** Version Date: 07.24.15		
The purpose of this checklist is to provide a set of specific elements necessary for a consistent and thorough data abstraction process for the Biosketches and Curriculum Vitae of key staff members from the NHLBI Proteomics Centers. The checklist is meant to serve as a quality review tool to be completed by an external reviewer during the interview process.		
1.0	**FILE STRUCTURE**	
☐	1.1	Abstractor has created a file according to the personnel's identification number.
☐	1.2	Abstractor has re-named and named files for that respondent according to the assigned identification number.
2.0	**ABSTRACTION CONDUCT**	
☐	2.1	Abstractor has filled in the header information regarding the abstraction.
☐	2.2	Abstractor has filled in the appropriate Position Advancement information (only if between 2010-2015).
☐	2.3	Abstractor has filled in the appropriate Honors and Service information (only if all or a portion of service fell on dates between 2010-2015).
☐	2.4	Abstractor has filled in the appropriate ongoing Non-NIH Research Support Funding information (only if between 2010-2015).
☐	2.5	Abstractor has entered the amount of ongoing funding in US Dollars.
☐	2.6	Abstractor has filled in the appropriate completed Non-NIH Research Support Funding information (only if between 2010-2015).
☐	2.7	Abstractor has entered the amount of completed funding in US Dollars.
☐	2.8	Abstractor has only entered N/A if information is not provided or there is no relevant information.
☐	2.9	Abstractor has completed the Protocol Memo noting methodological comments.
☐	2.10	Abstractor has completed the Protocol Memo noting data specific comments.
3.0	**NIH REPORT**	
☐	3.1	Abstractor has pulled Federal Funding into an excel file in the folder.

D.7. Biosketch/CV Protocol

Date	
Date of biosketch/CV Development (if available)	
File Type	
Center	
Person Code	

NHLBI BIOSKETCH/CV PROTOCOL – V 1.2

Element: Name

Position Title	**Department**	**Organization**	**Type Code:**	**Date of Advancement Between 2010-2015**
			ASSIST (1) - Assistant Professor	
			ASSOC (2) - Associate Professor	
			PRO (3) - Professor	
			NAMPRO (4) - Named Professor	
			DCHAR (5) - Department Chair	
			RES (6) - Researcher	
			SCI (7) - Scientist	
			ASSCI (8) - Associate Scientist	
			MENTOR (9) - Research Mentor	
			SCI (10) - Senior Scientist	
			ASDIR (11) - Associate Director	
			DIR (12) - Director or Co-Director	
			INT (13) - Intern	
			RES (14) - Resident	
			FELLOW (15) - Fellow	
			ASPHYS (16) - Assistant/ Associate Physician	
			PHYS (17) - Physician (Attending)	
			CHAR (18) - Chair/Chief	
			FAC (19) - Faculty Member	

Collect Advancements only within 2010-2015 – Typical Advancement Progress:

- Academic: Assistant Professor; Associate Professor; Professor; Distinguished (Endowed) Professor; Chair
- Research: Researcher; Scientist; Associate Scientist; Senior Scientist; Associate Director; Director (Co-)
- Medical: Intern; Resident; Fellow; Physician (Attending); Chair

Date	
Date of biosketch/CV Development (if available)	
File Type	
Center	
Person Code	

NHLBI BIOSKETCH/CV PROTOCOL – V 1.2

Element: Honors and Services

Full Name of Honor/Service	**Type of Honor/Service:**	**Date of Achievement/Service Between 2010-2015**	**Prior/Extending Achievement/Service Between 2010-2015**
	FEL (1) - Fellowship		
	LEC (2) - Lectureship		
	DPR (3) - Distinguished Prof.		
	AW (4) - Award		
	CMM (5) - Committee Member		
	CMC (6) - Committee Chair		
	POC (7) - Professional Org Member		
	POL (8) - Professional Org Leadership		
	EDB (9) - Editorial Board Member		
	EDL (10) - Editorial Board Leadership		
	RVC (11) - Review Committee Member		
	RVL (12) - Review Committee Leadership		
	SAB (13) - Science Advisory Board Member		
	SAL (14) - Science Advisory Board Leadership		
	PBR (15) - Publication Reviewer		
	INS (16) - Invited Speaker		

Date	
Date of biosketch/CV Development (if available)	

File Type	
Center	
Person Code	
Total # of Ongoing NON-NIH Research Support Fields	Insert # of Elements added below: #

NHLBI BIOSKETCH/CV PROTOCOL – V 1.2			
Element: Ongoing NON-NIH Research Support (Instructions: Place N/A in table sections where data cannot be found)			
Name of Project	**Acronym for Project**	**Agency Name**	**Acronym for Agency**
Type of Funding: NL (0) - Not Listed FD (1) - Foundational PR (2) - Private Company ST (3) - State UNV (4) - University	**Amount of Award** **US Dollars, Rounded**	**Timeframe** **0 = Overall** **1 = Current**	
Award Date: (mm/dd/yyyy)	**Award End Date: (mm/dd/yyyy)**	**Period of Funding:** W1 - Starting during 2010-2015 P2 - Stating prior to 2010-2015, but end data during 2010-2015	**Awardee's Role:** PRIN (1) - Principal investigator COPI (2) - Co-investigator/Joint PI CONS (3) - Consultant DIR (4) - Director MEN (5) - Mentor CHAIR (6) - Chairman MEM (7) - Member

Supporting Notes:

- If there is other 'Non-NIH Ongoing Research Support' copy and paste the above section to fill in below. If not, move on to Completed Research Support.
- If funding is in another currency other than US Dollars: use (http://www.oanda.com/currency/converter/) and enter the currency/award date to enter funding amount.

File Type	
Center	
Person Code	
Total # of Completed NON-NIH Research Support Fields	Insert # of Elements added below: #

NHLBI BIOSKETCH/CV PROTOCOL – V 1.2

Element: Completed NON-NIH Research Support (Instructions: Place N/A in table sections where data cannot be found)

Name of Project	**Acronym for Project**	**Agency Name**	**Acronym for Agency**

Type of Funding:	**Amount of Award US Dollars, Rounded**	**Timeframe 0 = Overall 1 = Current**
NL (0) - Not Listed		
FD (1) - Foundational		
PR (2) - Private Company		
ST (3) - State		
UNV (4) - University		

Award Date: (mm/dd/yyyy)	**Award End Date: (mm/dd/yyyy)**	**Period of Funding:** W1 - Starting during 2010-2015 P2 - Stating prior to 2010-2015, but end data during 2010-2015	**Awardee's Role:**
			PRIN (1) - Principal investigator
			COPI (2) - Co-investigator/Joint PI
			CONS (3) - Consultant
			DIR (4) - Director
			MEN (5) - Mentor
			CHAIR (6) - Chairman
			MEM (7) - Member

Supporting Notes:

- If there is other 'Non-NIH Completed Research Support' copy and paste the above section to fill in below.
- If funding is in another currency other than US Dollars: use (http://www.oanda.com/currency/converter/) and enter the currency/award date to enter funding amoun.t
- Click CTRL and A to select this whole document – choose font – CALIBRI
- Save documents to O:\Clients\NHLBI\G-Task 5. Data Abstraction\Biosketch and CV Completed Protocols
- Save this File to PERSON CODED FOLDER – as PERSON CODE BCV PROTOCOL.docx
- Save Biosketch/CV as PERSON CODE Biosketch.pdf or PERSON CODE CV.pdf

File Type	
Center	
Person Code	
NHLBI BIOSKETCH/CV PROTOCOL – V 1.2	
Element: NIH Funding	
Log-In to NIH RePORT and go to the QUERY page	
Enter in name of researcher on Biosketch/CV	
Go to Fiscal Year – check 2005-2015 – Click SELECT	
Scroll down and run query	
Click Export – Open Export file (.csv)	
Click Save As – Excel Workbook with Person Code FED FUNDING.xlsx in PERSON CODED FOLDER	

D.8. Biosketch/CV Codebook

25. Shared database	**SHARED_DATABASE**
Node	**Code**
Academic Advancement	
26. Faculty Member	FAC (19)
27. Assistant Professor	ASSIST (1)
28. Associate Professor	ASSOC (2)
29. Professor	PRO (3)
30. Named Professor	NAMPRO (4)
31. Department Chair	DCHAR (5)
Research Advancement	
32. Researcher	RES (6)
33. Scientist	SCI (7)
34. Associate Scientist	ASSCI (8)
35. Research Mentor	MENTOR (9)
36. Senior Scientist	SCI (10)
37. Associate Director	ASDIR (11)
38. Director or Co-Director	DIR (12)
Medical Advancement	
39. Intern	INT (13)
40. Resident	RES (14)
41. Fellow	FELLOW (15)
42. Assistant/Associate Physician	ASPHYS (16)
43. Physician (Attending)	PHYS (17)
44. Chair/Chief	CHAR (18)
Honors	
45. Fellowship	FEL (1)
46. Lectureship	LEC (2)
47. Distinguished Professor	DPR (3)
48. Award	AW (4)
Service	
49. Committee Member	CMM (5)

50. Committee Chair	CMC (6)
51. Professional Org Member	POC (7)
52. Professional Org Leadership	POL (8)
53. Editorial Board Member	EDB (9)
54. Editorial Board Leadership	EDL (10)
55. Review Committee Member	RVC (11)
56. Review Committee Leadership	RVL (12)
57. Science Advisory Board Member	SAB (13)
58. Science Advisory Board Leadership	SAL (14)
59. Publication Reviewer	PBR (15)
60. Invited Speaker	INS (16)
61. Not Listed/Undetermined/Multiple Funder Types	NL (0)
Pulling Information	
62. Foundation	FD (1)
63. Private Company	PR (2)
64. State	ST (3)
65. University	UNV (4)
66. Non-NIH U.S. Federal Agency Funding	UOFED (5)
67. International Government Agency Funding	IOFED (6)
68. Principal Investigator	PRIN (1)
69. Co-Investigator/Joint PI	COPI (2)
70. Consultant	CONS (3)
71. Director	DIR (4)
72. Mentor	MEN (5)
73. Chairman	CHAIR (6)
74. Member	MEM (7)