AAMC

Tomorrow's Doctors, Tomorrow's Cures®

# Institutional Approaches to Tracking Research Trainee Information

Association of
American Medical Colleges

## CONTENTS




This is a publication of the Association of American Medical Colleges. The AAMC serves and leads the academic medicine community to improve the health of all. **www.aamc.org**. Please address questions about the contents of this publication to Jodi Yellin, PhD, Director, Science Policy, AAMC (**jyellin@aamc.org**).

## EXECUTIVE SUMMARY

Information on career outcomes for biomedical PhD graduates and postdoctoral researchers is limited, and there is currently no comprehensive national system for tracking graduate students and postdoctoral researchers through their careers. The work described in this Association of American Medical Colleges (AAMC) report was undertaken to better understand what institutions and their PhD, MD/PhD, and postdoctoral programs are doing to collect research trainee information and to help institutions develop and enhance their own data-collection systems.

In 2012, a subgroup of the Group on Graduate Research, Education, and Training (GREAT Group) Steering Committee identified, through surveys and discussion, the types of data most institutions collect or would like to collect on their graduate students, postdoctoral researchers, and alumni. The subgroup then chose 12 representative AAMC-member institutions to participate in a study of their processes for collecting and disseminating information. The institutions reported using different strategies for data maintenance, and they varied in which trainees they collected data about and the type of software used. Respondents provided data on 20 databases in all.

Institutions identified which PhD, MD/PhD, and postdoctoral program fields were collected across five categories: program characteristics, faculty characteristics, incoming trainee populations, publications, career outcomes, and public information. Most databases collected program characteristics and incoming trainee population data, whereas other categories were less frequently tracked. Additional data acquired during interviews with seven of the institutions serve as the basis for the seven profiles in the report. The profiles include information about data-collection systems, data use, tracking career outcomes, and data maintenance.

The survey and interview data led to five major findings:

- Data-collection systems are varied.
- Automation and interoperability are primary technical challenges.
- Career outcomes data collection for all programs is incomplete.
- Postdoctoral research data are limited.
- Databases are used in multiple ways.

This report is intended to facilitate local and national discussions within the research and research trainee communities around research trainee data collection.

## INTRODUCTION

Biomedical PhD, MD/PhD, and postdoctoral trainees enter a wide range of careers in academia, government, industry, and other sectors. In 2012, the National Institutes of Health (NIH) Advisory Committee to the Director (ACD) Biomedical Research Workforce Working Group reported that 40 percent of biomedical PhDs now pursue careers in academic research or teaching, while others pursue careers in fields such as industry research, government research and administration, science writing, science policy, law, and consulting. The group also found that information on career outcomes for biomedical PhD graduates and postdoctoral researchers is limited, especially for individuals who earned their PhD degrees outside the United States. There is currently no comprehensive national system for tracking graduate students and postdoctoral researchers through their careers, although there have been numerous calls for one.[1, 2, 3]

The Association of American Medical Colleges (AAMC) and its Graduate Research, Education, and Training (GREAT) Group[4] have long been addressing and studying this important issue by providing a forum for sharing institutional practices, promoting awareness of multiple career pathways for biomedical research trainees, and providing an institutional perspective on workforce needs. The work described in this report was undertaken by the AAMC to better understand what institutions and their biomedical PhD, MD/PhD, and postdoctoral programs are currently doing to collect research trainee information and to help institutions develop and enhance their own data-collection systems.

Acquiring data about graduate and postdoctoral trainee outcomes is essential for institutional policy and decision making. These data provide key measures of success and a better understanding of the careers trainees enter, and they inform trainees of their career options as they plan to enter the biomedical workforce. At the national level, the research and research training community must ensure that we prepare the workforce to align with societal needs.

## METHODOLOGY

In 2012, the GREAT Group Steering Committee formed a subgroup to focus on understanding how institutions collect data on their graduate students, postdoctoral researchers, and alumni. The types of data elements most institutions collect or would like to collect were first identified through surveys and discussion. The subgroup then identified a set of 12 AAMC-member institutions to participate in a more in-depth study of their institutions' processes and procedures for collecting and disseminating information. The institutions had at least one trainee database and varied in terms of type of institution (private or public), size, and geographic location. A point of contact, whose role is to direct and/or administer at least one research training program, was identified for each institution. AAMC staff administered an online questionnaire to each institution, and after a preliminary analysis by AAMC staff and the GREAT Group data subgroup, seven institutional representatives were identified for follow-up interviews. These discussions addressed more in-depth questions about the institutional database(s) and were used as the basis for the profiles in this report. The data collection in this study was conducted according to AAMC data policies and procedures.

## RESULTS

### Data-collection systems

Twelve AAMC-member institutions, eight private and four public, were selected to participate in this study. The number of current PhD, MD/PhD, and postdoctoral biomedical research trainees reported by each institution ranged from about 300 to nearly 2,000. Five of the institutions reported that their databases covered training programs in the medical school only, four were university-wide, and three covered specific programs within the graduate school and/or the school of medicine.

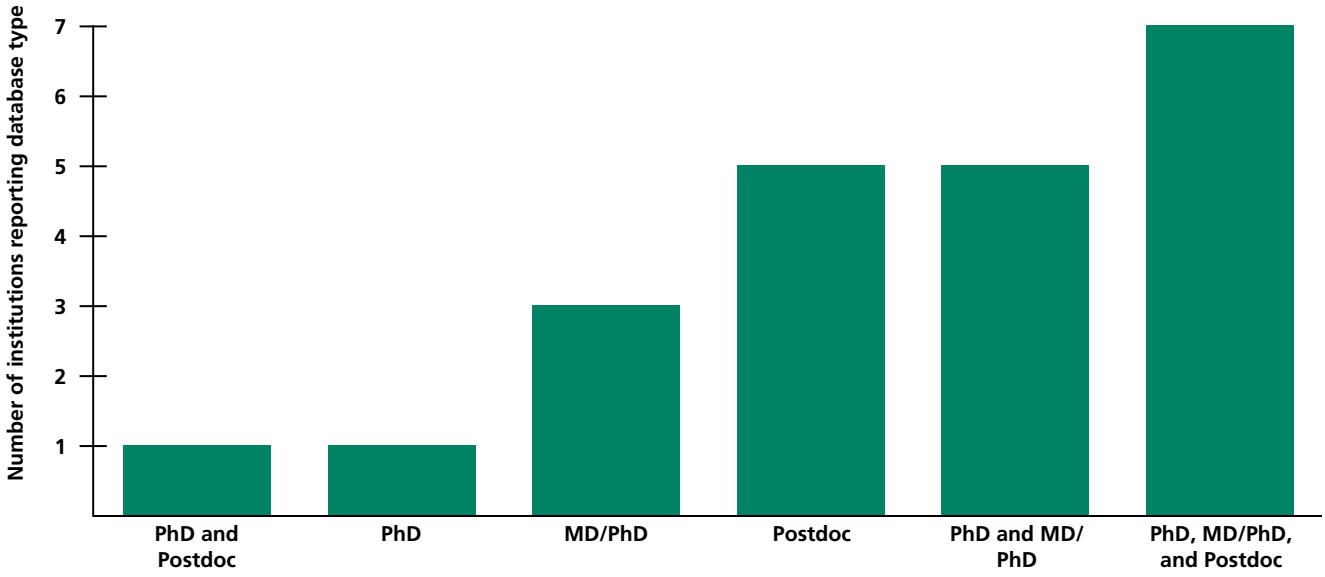
The 12 institutions reported using different strategies for data maintenance. For example, seven had a single database that contained information on PhD, MD/PhD, and postdoctoral researcher trainees (**Figure 1**). Two of these seven also had separate databases for different trainee groups that were not coordinated with each other. The other institutions either reported using different databases to track different groups of trainees or did not collect data on one or more trainee groups. In some cases, PhD and postdoctoral researcher data resided within the same database, and in others, PhD and MD/PhD data were combined. Questionnaire respondents subsequently provided data on 20 of the 22 reported databases.

Database development processes differed widely among the institutions. Indeed, many respondents were not able to identify the length of the development process for their database(s). Of the institutions that provided a time frame, the majority of estimates ranged from one to six years. During the interviews, many noted that institutional programs went through different versions to reach their current database and are still revising their systems as needed, implementing improvements such as adding fields, streamlining organization, and increasing interoperability with other campus systems. Eighty percent of the databases were built using commercial software, including Microsoft Access, FileMaker Pro, Curvita, Banner, Saleslogix, PeopleSoft, and Oracle. Twenty percent of the databases use completely homegrown systems, and 70 percent use some type of institutionally developed software to assist with uploading or updating data, even when using a commercial system.

### Database fields

Institutions were asked to identify which PhD, MD/PhD, and postdoctoral program fields were collected out of a list of preidentified fields broken out by category: program characteristics, faculty characteristics, incoming trainee populations, publications, career outcomes, and public information (see **Figure 2**, on pages 22–24).

**Figure 1. Types of trainee databases reported by the 12 institutions queried.**



The program characteristics and incoming trainee population fields were most commonly collected for PhD, MD/PhD, and postdoctoral programs (Figures **2A** and **2C**). Although most institutions collected the number and gender of faculty appointed to programs, the majority of institutions did not collect other faculty-characteristic fields (**Figure 2B**). Publication data were collected for 7 out of 11 institutions that collect PhD and MD/PhD program data, but these data were available for only 2 out of 10 institutions that collect postdoctoral program data (**Figure 2D**).

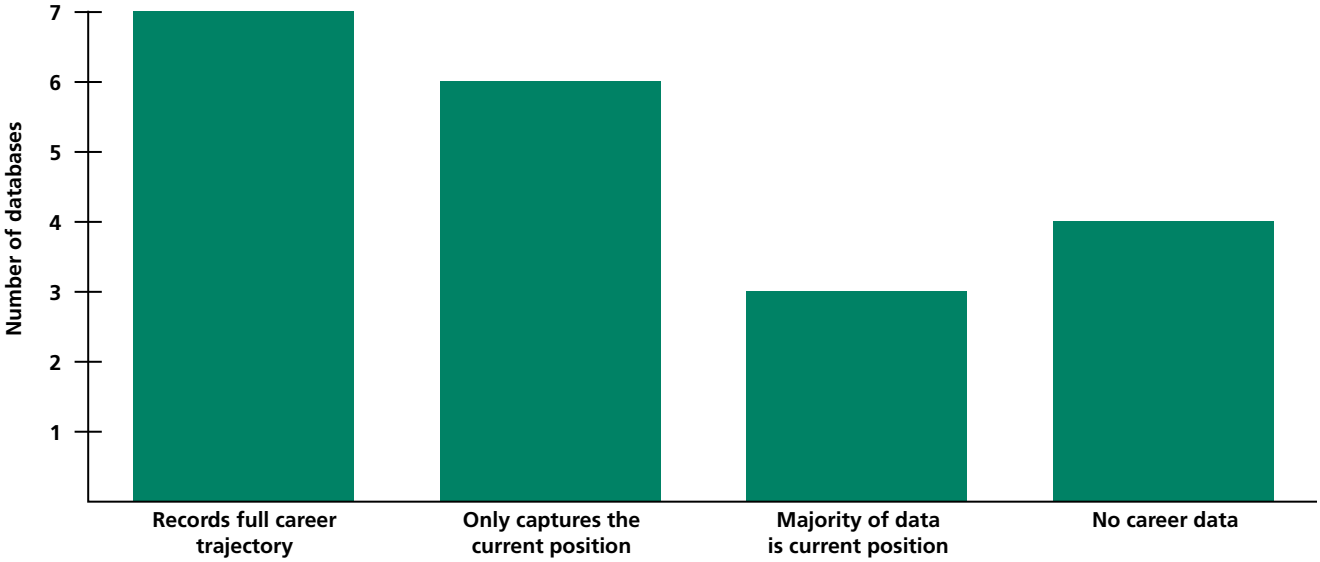
Approaches to collecting data on career outcomes varied across the institutions. While most institutions collect some career outcomes data for PhD and MD/PhD programs, less than half collect career outcomes data for postdoctoral programs (**Figure 2E**). In most cases, institutions collect information about current positions; however, only 35 percent of the databases keep a full record of career trajectory (**Figure 3**). The percentage of prior trainees for whom any career outcome data were available varied widely between 30 and 100 percent. How long individuals are tracked also varied widely, with some respondents collecting data for three years and others having no limit, that is, tracking for as long as possible.

Of the 20 different databases analyzed in the study, 14 have some data fields that are available to the public on the program website—for example, in an annual report or as responses to individual data requests. Four of the six databases that do not make any data public are postdoctoral-only databases, with one institution commenting that "very little is known about postdoctoral career trajectories." In addition, just 2 out of the 10 databases that include postdoctoral program data report public information on postdoctoral researchers. Only 1 institution makes public career outcomes data for postdoctoral researchers, compared with 7 out of 11 institutions that collect PhD program data and 7 out of 11 institutions that collect MD/PhD program data (**Figure 2F**).

## INSTITUTION PROFILES

The following profiles highlight some of the key components and strategies of the major institutional database at seven different institutions.

**Figure 3. Career trajectory information captured by institutional databases.**




AAMC

# Emory University

## DATABASE OVERVIEW

- **Database established:** 1995
- **Who authorized database creation:** Director, Graduate Division of Biological and Biomedical Sciences
- **Trainees covered by database:** PhD, MD/PhD, Postdoctoral
- **Database software:** FileMaker Pro
- **Manual or automatic data entry:** Both
- **Number of current trainees:** ≈440 PhD and MD/PhD, 800 Postdoctoral

## I. Data-Collection System

### System Establishment and Ownership

The database is managed by the Emory Graduate School and includes PhD students in the biological and biomedical science programs and postdoctoral researchers in the School of Medicine.

### Driver for System Creation

The database originally was set up to gather information for training grants but has evolved for other purposes, such as recruitment and program quality evaluation. The database was first built for tracking PhD students, was expanded two years ago to include postdoctoral researchers, and is currently being expanded to include MD/PhD students.

### Software

The database is stored in FileMaker Pro and Oracle.

## FOCUS: REDESIGN TO A RELATIONAL DATABASE

The database originally was created as a flat database, designed around a single table and without automation in data sharing. Two years ago, the database was redesigned as a relational database with multiple tables that link to other institutional databases. This allows for better integration and collation of the data.

## II. Data Use

### Program and Institutional Use

***Recruitment:*** An anonymized list of student outcomes is available on the Emory website, listing trainees' graduation year, program, and current position. This provides a sense of possible career paths for prospective applicants and student advisors. Postdoctoral outcomes are not currently collected in the Graduate School database but will be added to the Oracle postdoctoral researcher database that resides in the Office of Postdoctoral Education in the School of Medicine.

***Program Evaluation:*** Alumni outcomes are routinely assessed during program evaluation and university administration review.

***Planning and Development:*** The programs do not directly solicit donations but do share information with the university alumni development office, which tracks both graduate students and postdoctoral researchers. The development office is also establishing endowed student awards, for which a record of alumni is useful.

### Data Sharing and Privacy Considerations

The data are used for reporting to federal funding agencies, including the National Institutes of Health (NIH) and the National Science Foundation (NSF). They are also used during the grant application process and to respond to queries and polls from colleagues. The systems are behind a firewall and university-level security, and data are scrubbed of personal identifiers such as Social Security numbers.

## III. Tracking Career Outcomes

### Database Coverage

The database covers about 60 percent of alumni, though an aspirational rate of coverage is 80 to 90 percent for graduate students and 70 to 80 percent for postdoctoral researchers. Trainees are tracked for as long as possible, and this process is facilitated when contact with alumni begins immediately following graduation. This enables tracking for each position as it occurs in the trainee's career. It is more difficult to track alumni 15 to 20 years out, though the quality of the data isn't compromised in the cases where the data are obtainable.

### Data-Gathering Strategies

Alumni prefer to be contacted via email, although LinkedIn is frequently used to find career information. It can be difficult to obtain current alumni contact information, but once that is resolved, alumni are generally very responsive to data requests.

## FOCUS: CAREER BUILDING FOR TRAINEES

If a student expresses interest in a particular career, such as a job in a certain sector or with a specific federal agency, the program office attempts to facilitate a connection or set up an informational interview with relevant alumni by using the career fields in the database.

*Continued on next page*




# Emory University *(continued)*

## IV. Data Maintenance

### Keeping Data Current

The initial data for a student's record flow from the admission application, and these data are subsequently updated throughout the trainee's lifespan and once or twice a year for alumni. For postdoctoral researchers, the data flow from initial registration and human resources records at the time of new postdoctoral researcher orientation sessions and are supplemented with Web-based forms for completing individual development plans, annual benchmark progress reports, and exit interviews. If possible, the data are pulled automatically—for example, when adding transcripts from the registrar system or postdoctoral researcher information from PeopleSoft, the primary university database. However, updates to fields such as advancement to candidacy are still manually entered into the database through Web-based interfaces. Current students, postdoctoral researchers, alumni, and faculty can update records directly and enter publications, awards, honors, and other relevant information into the database. Having a Web interface where individuals can correct their own information has been key for maintaining an updated database.

### Technical Challenges and Future Needs

One technical challenge has been pulling data from the Oracle-based PeopleSoft database into the FileMaker Pro database because it requires writing programs to interface between the two systems. The database is currently undergoing a major rewrite to use a more recent version of FileMaker Pro. Information needed to complete training grant tables is compiled in Oracle reports. The Graduate School is looking into building a similar database for departments and disciplines beyond the biomedical sciences. A future goal is to use the database to better identify applicant predictors of career success.

### Personnel and Financial Support

An outside consultant, funded by the Graduate School, was hired as a FileMaker developer to set up the system. University Information Technology (IT) resources were used when the databases were expanded to include postdoctoral researchers and MD/PhDs. Depending on the type of trainee, the database is supported by funds from either the Graduate School or the School of Medicine. Currently, a full-time database manager is responsible for data entry and ensuring data integrity. The database manager also generates the training tables for institutional NIH T32 applications. Support staff are in each of the individual graduate programs and in the postdoctoral office.


# Tufts University Sackler School of Graduate Biomedical Sciences

## DATABASE OVERVIEW

- **Database established:** 2003
- **Who authorized database creation:** Dean of the Graduate School
- **Trainees covered by database:** PhD, MD/PhD, Postdoctoral
- **Database software:** Microsoft Access
- **Manual or automatic data entry:** Manual
- **Number of current trainees:** ≈260 PhD and MD/PhD, 200 Postdoctoral

### I. Data-Collection System

**System Establishment and Ownership**
The system was established by the dean of the Graduate School and currently operates under the Graduate School.

**Driver for System Creation**
The database was set up so the institution can understand the career outcomes of program alumni and provide information for training grant reporting.

**Software**
The database is stored in Microsoft Access.

### II. Data Use

**Program and Institutional Use**
***Recruitment:*** Data on outcomes (primarily PhD and MD/PhD) are published on the institution's website and are available for potential or current trainees.
***Program Evaluation:*** Career outcomes are tracked to ensure that programs provide training for the positions that alumni are pursuing. Fields such as time to degree, competitive fellowships, and publications are tracked as part of the program review process; diversity and disability data for PhD and MD/PhD students are also included. In addition, data are used at the university-wide level to advocate for the importance of graduate education.
***Planning and Development:*** Some of the data are shared with the development office, which takes the lead on any fundraising activity.

**Data Sharing and Privacy Considerations**
The data are shared for grant applications (federal and nonfederal) and for institutional accreditation. Data accessibility is limited and is subject to the same protections (e.g., de-identification) as are all other trainee or employee data.

### III. Tracking Career Outcomes

**Database Coverage**
The database covers 100 percent of all PhD and MD/PhD alumni and close to 90 percent of postdoctoral researchers.

**Data-Gathering Strategies**
Trainees are tracked via email, LinkedIn, and Facebook (which is used primarily to find updated contact information). A small amount of this communication also goes out through the university development office.

## FOCUS: STRATEGIES FOR TRAINEE ENGAGEMENT

Having trainees connect with the institution before they finish their program facilitates long-term contact, as does creating a culture where multiple career choices are valued. It has been found that trainees are more likely to be open and willing to share information in an accepting environment.

### IV. Data Maintenance

**Keeping Data Current**
All alumni have been tracked since the establishment of the program (40 years ago). The data are updated every four to six months or more frequently if the school is notified of new information by an alum.

**Technical Challenges and Future Needs**
The database does not currently interface with other university systems. In the future, the data may be moved from Microsoft Access to FileMaker Pro, which would facilitate automatic data transfer. Ideally, with additional resources, the database would be expanded to include data on career trajectories and more current postdoctoral data.

**Personnel and Financial Support**
Funding for the database is from the Graduate School, which operates under the School of Medicine budget with support from the Office of the Provost. The data are managed by personnel in the graduate office.

## FOCUS: BUILDING A DATABASE TO SERVE THE NEEDS OF THE INSTITUTION

As a smaller school tracking a few thousand current and former trainees, the necessary data are obtained and stored using easily available software and minimal personnel. It has been helpful to the institution to create a balanced system that meets local needs and measures program performance without an undue strain on limited resources.


# University of Alabama at Birmingham, Graduate Biomedical Sciences

## DATABASE OVERVIEW

- **Database established:** 2008
- **Who authorized database creation:** Dean of the Graduate School; Provost
- **Trainees covered by database:** PhD, MD/PhD
- **Database software:** Excel
- **Manual or automatic data entry:** Both
- **Number of current trainees:** ≈370 PhD and MD/PhD, 230 Postdoctoral

## I. Data-Collection System

### System Establishment and Ownership

The development of the system was initiated by the senior associate dean of the Graduate School, and the database operates under the Graduate Biomedical Sciences Division.

### Driver for System Creation

The overall purpose of the system is both to track both career outcomes and to support training grant preparation. The data collected include details of the application and recruitment process, coursework, degree milestones, funding, publications, and career outcomes.

### Software

The PhD-MD/PhD database is in Excel. (It was initially set up using FileMaker Pro, but this software is no longer used.) UAB postdoctoral researcher data are currently being moved from a paper-based to an electronic system.

## FOCUS: EXTERNAL REVIEW

A review of the Graduate School by outside consultants (Huron Consulting Group) recommended improvements to the database infrastructure, which led to the approval of additional funds to enhance data access.

## II. Data Use

### Program and Institutional Use

***Recruitment:*** The Biomedical Sciences Division tracks applications, interviews, and graduate program offers. The profile of the incoming class and alumni career outcomes are used for advertising purposes.

***Program Evaluation:*** Data are used internally for institutional program review and an annual PhD program review and for external purposes, such as LCME review or NIH training grant reports.

***Planning and Development:*** Data are provided to the alumni office for fundraising campaigns. Information is also shared with the university development office, which has been useful for soliciting contributions. Data from the biomedical sciences are also combined with a larger data set to determine funding allocation for first-year Graduate School fellowships.

### Data Sharing and Privacy Considerations

The aggregate data are available on various program websites and in other printed materials. The data are protected and available through an institutional SharePoint site, and only designated individuals are allowed to modify the data.

## III. Tracking Career Outcomes

### Database Coverage

The database collects information on trainee career outcomes and trajectory for as long as possible after the graduation date. For PhD trainees, the database covers about 60 percent of alumni, and, ideally, it would reach a benchmark of at least 70 percent to provide statistically relevant information on outcomes. The MD/PhD program aims to track 100 percent of its graduates, and it currently has information on all their alumni. This is likely because the career paths of MD/PhDs after graduation are less varied than those of PhDs.

### Data-Gathering Strategies

Information is gathered through direct emails to graduates and by an alumni tracking service, inDegree. Alumni are contacted six months after graduation to confirm their first position, and inDegree is used every other year to ensure information is fully updated. If there is a particular request from another university entity on career outcomes for a selected group or certain individuals, the office may pull that data manually from Web searches or any other available sources.

## FOCUS: inDegree

inDegree, an information-gathering tool developed at UAB's business school, mines the self-reported data on LinkedIn to collect and maintain accurate data on alumni careers. These data are combined with alumni data from university registration systems to create a full dataset.
**www.indegree.com**

*Continued on next page*


# University of Alabama at Birmingham, Graduate Biomedical Sciences *(continued)*

## IV. Data Maintenance

**Keeping Data Current**

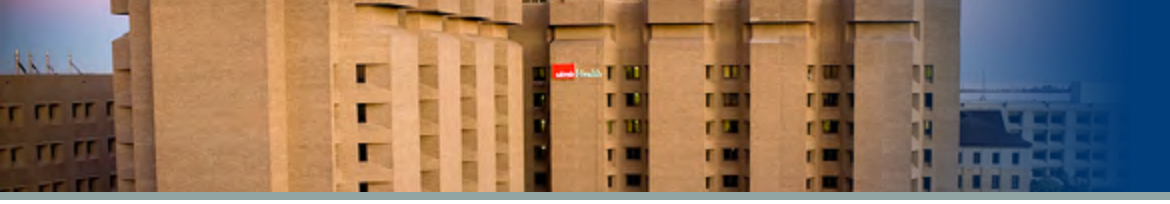
Active student and applicant data are automatically updated overnight. The database partially stems from student information stored on Banner, and it also pulls data from Oracle (the university HR system) and the AMCAS® medical school application for MD/PhD students. One of the challenges in linking with university systems has been obtaining the necessary authorization and access to pull information into the local trainee tracking system. A few months in advance of training grant deadlines, the program reviews the information in the database on current students and alumni to ensure that it is as complete as possible.

**Technical Challenges and Future Needs**

Currently, as application information is updated for a given applicant, there is a significant amount of daily reviewing and updating; the program office would like to automate this process via new homegrown software. UAB staff are working with university IT to develop in-house means to automate data recovery and management from disparate sources, particularly because manual data input by university staff has the potential to lead to errors or an inaccurate assessment of a data-field value. The program office is also reformatting NIH T32 training tables and developing a Web-based form for updating alumni content. Finally, the program would like to develop an accurate and effective method for collecting publications and current email addresses from trainees.

**Personnel and Financial Support**

The Graduate School supports the everyday infrastructure costs and collaborates with the postdoctoral office to fund the data support specialist, who manages databases for both pre- and postdoctoral researchers.



# University of Texas Southwestern Medical Center

## DATABASE OVERVIEW

- **Database established:** 2009
- **Who authorized database creation:** Vice Provost
- **Trainees covered by database:** PhD, MD/PhD, Postdoctoral
- **Database software:** In-house
- **Manual or automatic data entry:** Both
- **Number of current trainees:** ≈450 PhD and MD/PhD, 600 Postdoctoral

### I. Data-Collection System

**System Establishment and Ownership**
The vice provost authorized the establishment of the data-collection system, and the associate dean of the Graduate School directed its development. The system is owned by the Graduate School.

**Driver for System Creation**
The driving force behind setting up a robust tracking system was to create a system that could be used (1) as a recruiting tool, (2) for potential fundraising, (3) to maintain a relationship with alumni, and (4) for data collection and reporting.

**Software**
The trainee tracking system is composed of a SQL Server database with a Web user interface. UT Southwestern developed the homegrown database with resources provided by the university's academic information resources team. The most recent version of the database launched in September 2014.

### II. Data Use

**Program and Institutional Use**
***Recruitment:*** General trainee career outcomes are presented to PhD program applicants. The MD/PhD program has a much smaller population and uses outcomes data for potential applicants more heavily. Postdoctoral outcomes data are not currently shared. One of the goals of the new database is to be able to provide more detailed outcomes data to applicants. Another goal is to engage alumni to act as recruiters for the PhD and MD/PhD programs.
***Program Evaluation:*** Centralization of the data collected and reported by the Graduate School for the purposes of NIH training grant applications and other surveys improves the effectiveness of program evaluation.
***Planning and Development:*** Donor relations are managed currently by the university development office. However, in the future, data may be used to interest alumni in investing in initiatives specific to the Graduate School. Increased relationships with alumni will also provide networking and speaker resources for the Graduate Career Services Office.

**Data Sharing and Privacy Considerations**
The data are shared outside the institution for training grant applications, state reporting (e.g., the Texas Higher Education Coordinating Board), accreditation purposes, and any other surveys for federal agencies or nonprofits.

### III. Tracking Career Outcomes

**Database Coverage**
The database covers about 50 percent of PhD alumni, 95 percent of MD/PhD alumni, and 50 percent of postdoctoral researcher alumni. The aspirational rate of coverage for the database is about 90 percent. The database was set up initially to include alumni 15 years past graduation. The interest in tracking alumni after that time period does not diminish, but it does become a less feasible process. One issue has been having enough staff to track alumni as they get further past graduation, particularly because this tends to be a manual process. The goal is to add every alumn of a UT Southwestern program and to be in touch with them for the rest of their career.

**Data-Gathering Strategies**
Multiple strategies are used to gather data, including reaching out through the trainee's mentor, emailing alumni directly, and looking up information through social media, primarily LinkedIn. The information is validated by cross-checking these various sources. Current alumni contact information is sometimes found through their connection with other alumni.

## FOCUS: USING DATA FOR ONGOING COMMUNICATION AND DISSEMINATION

A future plan is to use the data to reach out to alumni semiannually with program updates to (1) develop strong relationships between alumni and the graduate enterprise and (2) enable those alumni who mentor trainees to encourage their current trainees to participate in undergraduate research and educational opportunities at UT Southwestern.

*Continued on next page*

# University of Texas Southwestern Medical Center *(continued)*

## IV. Data Maintenance

**Keeping Data Current**
When building the database, an attempt was made to automate as many fields as possible. Data are collected on a continuous basis, but how often they are updated in the database depends on staff availability.

**Technical Challenges and Future Needs**
One of the major challenges has been ensuring compatibility with other university systems; when they are changed, the database also needs to be adjusted to continue to interface with institutional data. Another technical issue has been interfacing with the PeopleSoft software used by the university. Possibilities for the database in the future include adding a numerical identifier (e.g., NIH Commons or ORCID ID) and exploring options to mine publications from PubMed.

**Personnel and Financial Support**
Currently, the associate dean has protected time (10 percent) to oversee the development of a separate alumni system, in addition to a database for current trainees. There are plans to have a full-time equivalent employee, funded through the Graduate School, populate and maintain the data and continue software support from the university's Academic Information Resources Office.

**FOCUS: COLLATING DATA ACROSS TRAINEES**
One of the features of the new database is the ability to look at how many students and postdoctoral researchers a given faculty member has trained over the past 15 years, which will help provide insight into faculty mentoring.

# University of Virginia School of Medicine

## DATABASE OVERVIEW

- **Database established:** 1988
- **Who authorized database creation:** Assistant Dean for Graduate Research and Training, School of Medicine
- **Trainees covered by database:** PhD, MD/PhD
- **Database software:** Curvita-SciMed Solutions
- **Manual or automatic data entry:** Both
- **Number of current trainees:** ≈400 PhD and MD/PhD

## I. Data-Collection System

### System Establishment and Ownership

The system operates under the School of Medicine.

### Driver for System Creation

The initial goal of the system was to generate a standard set of NIH training grant tables. Data were first collected for one department, and the database was expanded in 1996 to serve all graduate programs. It is customized to the needs of the institution; all fields needed for the current NIH training grant tables are collected.

### Software

The database was originally built using Microsoft Access but transitioned to commercial software known as Curvita. This migration allowed for greater flexibility in programming and extracting data. Curvita is no longer commercially available because of the cost and burden of creating software able to interface with different university systems at individual institutions.

## FOCUS: DATABASE AS AN ADMINISTRATIVE TOOL

Information from the database, such as publication records and faculty availability to take on new trainees, is exported automatically to about a dozen different campus webpages. This decreases the administrative burden on faculty of updating information in several different places.

## II. Data Use

### Program and Institutional Use

***Recruitment:*** Graduation rates and data on student outcomes are made available to program applicants.
***Program Evaluation:*** The university conducts a review of all graduate programs on a five-year rolling basis, which includes a review of data on graduation and attrition rates from the database.
***Planning and Development:*** The database covers all graduate programs that are affiliated with the biomedical sciences as well as the School of Medicine, which allows for the data to be used to give an overall financial picture of how funding flows between these different parts of the institution. The data are not used for fundraising purposes.

### Data Sharing and Privacy Considerations

Data are shared outside the institution for accreditation purposes. When data are made available, they are de-identified and usually provided in table format to discourage manipulation. Occasionally, specific fields are shared in response to an individual request. Administrators and faculty are permitted access to the nonconfidential information in the database but are not able to modify or delete data. There are multiple levels of security at the university level, including storage of the data on secure nonlocal servers, which require authentication for access.

## FOCUS: QUICK REACTION TO FUNDING OPPORTUNITIES

When a new NIH T32 training program in biodefense was introduced, UVA was able to submit a grant application within two months of the funding opportunity announcement by quickly compiling a list of existing biodefense-related projects via a database search of publications and dissertation titles. The institution was funded in the first round of awards.

## III. Tracking Career Outcomes

### Database Coverage

The database keeps a record of career trajectory for as long as possible, but it becomes more difficult to obtain accurate data the longer trainees have been out of the program. Program administrators are most confident about the quality of the data from the 10-year period following graduation, which is also the time frame required for training grant reporting. The success rate for collecting alumni information was previously around 65 percent, but it has dropped over the past few years as a result of budget-related staff cuts and less time to verify data. However, the trend in data collection suggests there will be more success in tracking students going forward than there has been in the past because of the use of electronic communication and Internet resources.

*Continued on next page*

# University of Virginia School of Medicine *(continued)*

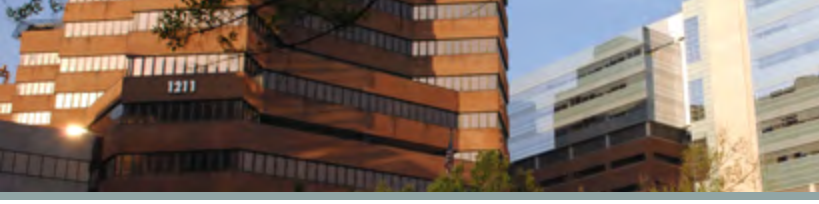
**Data-Gathering Strategies**
Current students are likely to join the program's LinkedIn group before they graduate, which allows for automatic and effective tracking after they leave the institution. Program staff will also periodically send an email or do a Web search to obtain updated information, but they have found that trainees prefer LinkedIn to other forms of correspondence.

## IV. Data Maintenance

**Keeping Data Current**
Information can be exported from, but not imported to, the database. In order to incorporate data from other university systems, such as Oracle, the information is pulled via a query and then manually formatted and entered into the trainee database. Faculty can directly enter relevant information, such as updates to their students' publications or new positions. Graduate administrators also make changes to the database as needed. Occasionally, these updates are spot-checked to verify data accuracy. The goal is to have multiple inputs into the database to ensure that the information is as complete and up to date as possible.

**Technical Challenges and Future Needs**
The existing software fulfills the current needs of the program. A desired functionality for the future would be the ability to import funding data from the NIH, but this is extremely difficult to do accurately because of multiple start dates and funding cycles.

**Personnel and Financial Support**
The database is funded by the School of Medicine, under the interdisciplinary Graduate Programs Office. Other departments throughout the institution have also expressed interest in this type of database, and there has been investment from the provost's office via staff support to build that capacity. It is possible that some form of the database will shift to the university level in the future.



# Vanderbilt University Medical School

## DATABASE OVERVIEW

- **Database established:** 1999
- **Who authorized database creation:** Senior Associate Dean, School of Medicine
- **Trainees covered by database:** PhD, MD/PhD, Postdoctoral
- **Database software:** In-house
- **Manual or automatic data entry:** Both
- **Number of current trainees:** ≈640 PhD and MD/PhD, 520 Postdoctoral

## I. Data-Collection System

### System Establishment and Ownership

The senior associate dean for biomedical research, education, and training in the School of Medicine was responsible for the establishment and development of the system. He now has the authority to access the full database and is accountable for making any key database-related decisions.

### Driver for System Creation

The system was created to track the life cycle of a student through the graduate program, from recruitment to training and beyond, as well as to assist with training grant preparation.

### Software

The software was developed in-house. The database has undergone three redesigns since its establishment.

## FOCUS: NIH BEST AWARD

The institution is a recipient of the NIH Broadening Experiences in Scientific Training (BEST) Award for biomedical research workforce innovation. Success in receiving the BEST Award, training grants, and other awards helps secure continued resource support from the institution for the database.

## II. Data Use

### Program and Institutional Use

***Recruitment:*** Data on trainee retention, demographics, and outcomes are provided to potential applicants and incoming students.

***Program Evaluation:*** Students complete evaluations during their first and final year that are entered into the database, which allows the institution to track satisfaction, productivity, and the student's development as a scientist. Individual programs are also assessed based on these evaluation data and other student-acquired data fields, such as number of publications and postgraduate career outcomes. Postdoctoral researcher data are also used to evaluate funding and career outcomes.

***Planning and Development:*** The database is used to assist in financial planning by collecting information on training grants and faculty awards and developing an overall picture of the institutional resources available to support research training. These data are also used to determine program size.

### Data Sharing and Privacy Considerations

The data are shared in NIH training grant applications and are provided to NSF in response to surveys on graduate students and postdoctoral researchers. They are also shared with other institutions for learning purposes when requested.

## III. Tracking Career Outcomes

### Database Coverage

The first trainees graduated from the umbrella PhD program in 1994; currently, the database tracks trainees for up to 20 years. The completeness of the alumni data is partially based on the ease of finding information and may not reach 100 percent coverage because of the difficulty in tracking some trainees.

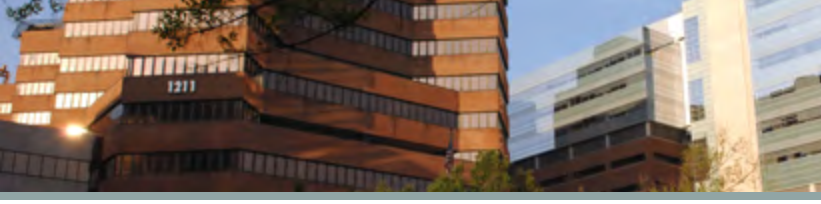
### Data-Gathering Strategies

Data are gathered from multiple sources (i.e., phone calls, websites, social media) to ensure database accuracy. High school students are hired periodically to assist in data gathering. The program uses this method to track up to 85 percent of alumni.

## FOCUS: DATA-GATHERING STRATEGIES

The program periodically hires a high school student to find information on alumni by doing Web searches and gathering data from social media.

*Continued on next page*



# Vanderbilt University Medical School *(continued)*

## IV. Data Maintenance

**Keeping Data Current**
The software interfaces with other university systems, including the registrar's, admissions', and HR's. Faculty data are updated as needed for competitive grant renewals. Data updates are automated whenever possible; some data, such as rotation evaluations, are entered manually.

**Technical Challenges and Future Needs**
Technical issues are generally solved by a software developer, though significant challenges can arise when other entities within the university change to new software. Other major challenges lie in collaborating with different parts of the university to have access to the necessary data and in having continuity in the personnel who work with these data. Financial stress on academic medical centers has exacerbated some of these problems.

**Personnel and Financial Support**
The database was built initially by an IT team at the university, but the design, structure, and data entry are run through the Office of Biomedical Research Education and Training (BRET). Since 2011, IT staff support has been centralized at the institution level. Resources for system development were provided by the School of Medicine. Several staff from the BRET office spend a portion of their time on database maintenance and work with the university IT team if any design updates are needed.

# Yale University, Biological and Biomedical Sciences Program

## DATABASE OVERVIEW

- **Database established:** 2014
- **Who authorized database creation:** Medical School Deputy Dean for Finance
- **Trainees covered by database:** Postbaccalaureate, PhD, MD, MD/PhD, Postdoctoral
- **Database software:** FileMaker 13 Pro
- **Manual or automatic data entry:** Both
- **Number of current trainees:** ≈580 PhD and MD/PhD, 1,300 Postdoctoral

## I. Data-Collection System

### System Establishment and Ownership

The system originated in and is owned by the Medical School. The deputy dean for finance authorized the establishment of the database, and the IT director of health and medicine is driving system development.

### Driver for System Creation

The main objective of the database is to assist in effectively preparing predoctoral training grant applications. Every field that the NIH currently requires for a training grant submission is being collected.

### Software

The data are stored in FileMaker Pro, which allows programmers to pull data from other university systems. The decision was made to use off-the-shelf software and then modify it specifically for database needs because of the lack of commercial software that allowed interaction with preexisting campus databases.

## II. Data Use

### Program and Institutional Use

***Recruitment:*** Once the database is more complete, data may be used to highlight accomplishments of current and past trainees.

***Program Evaluation:*** The database will be used to populate and generate NIH T32 tables, initially for predoctoral training grants and eventually for postdoctoral grants. The data collected will also enable the MD, PhD, and MD/PhD programs to analyze data points, such as time to degree and the number of publications per student. Individual student progress also will be tracked.

***Planning and Development:*** A planned use of the data is for financial modeling; funding sources for each trainee will be tracked with the goal of making projections to help plan for future years.

### Data Sharing and Privacy Considerations

Data are stored behind a firewall, and only a limited number of users will have access to the database.

## FOCUS: INTERNAL REPORTS

A goal is to build the capacity for various MD and PhD programs to use the database to generate reports on their trainees and explore data points, such as time to degree or metrics for an individual investigator or lab. This would allow departments to query the data according to their specific needs.

## III. Tracking Career Outcomes

### Database Coverage

Because the database is new, only a small percentage of trainees are currently tracked. The goal is to track every bioscience trainee, whether or not that individual was ever on a training grant. A realistic estimate for coverage is 80 percent, though an aspirational rate would be all trainees.

### Data-Gathering Strategies

The current strategy is to manually gather career outcomes data by contacting individual trainees and their former faculty mentors and searching LinkedIn. This information can be difficult to validate, so future plans are to use internally developed software to automatically contact alumni in order to incorporate updated information directly into the database.

*Continued on next page*

# Yale University, Biological and Biomedical Sciences Program *(continued)*

## IV. Data Maintenance

**Keeping Data Current**
Data were initially entered into the system manually from previous training grant applications, faculty input, and reports from faculty, student, and financial databases. There is currently a plan to implement quarterly updates from the university grants and contracts software system. Alumni data will be updated yearly. Additionally, the assignment and use of ORCID identifiers will be used to retrieve applicable publications from PubMed.

**Technical Challenges**
Current technical considerations include deciding who will have access to the data to modify it and how to reconcile data in different formats from across university systems. Working with data and records has presented difficulties in the past when there was no standardized method for entry (e.g., how first and last names are listed in a database).

**Personnel and Financial Support**
The database is financially supported through the Medical School.

**FOCUS: ADAPTABILITY TO FUTURE NEEDS**
Ensure that the database can be easily modified or adapted for future needs, such as changed requirements for training grant tables.

## DISCUSSION

**Data-collection systems are varied.**
While all institutions participating in the study recognized the value of establishing and maintaining a database, their approaches to collecting and storing data on their research trainees and program alumni vary. Database ownership and authorization are primarily at the level of the provost or dean of the graduate or medical school. Financial support generally comes from within the budget of the graduate school or school of medicine. Other institutional resources in the form of expertise or personnel may also be used in system creation and maintenance. Many institutions also noted that existing staff within the graduate or postdoctoral program assist in database efforts. Engaging research and research training leaders in discussions at their institutions facilitates the establishment of shared goals for databases and helps identify the best approaches to achieve institutional goals within available resources.

**Automation and interoperability are primary technical challenges.**
Integrating a database with other university systems requires reconciling data in distinct formats, having access to data stored by different entities at an institution, and maintaining data privacy. The issue of data privacy is tackled at multiple levels, with institutions citing de-identification, limiting data access, and storage behind a firewall as potential solutions. For some databases, current trainee data are imported automatically from other university systems; however, much of the data updating for previous trainees continues to be done manually. Manual entry causes the databases to be more prone to errors. Several institutions cited collaboration across the institution as being important in overcoming technical challenges and ensuring successful data transfer.

**Career outcomes data collection for all programs is incomplete.**
In general, programs collect a limited amount of data on career outcomes for graduate students and postdoctoral researchers. Institutions noted the lack of standard definitions as well as limited staff time as barriers to consistent and complete data collection. Most institutions cited the advent of social media as critical to maintaining contact with and receiving updated data from former trainees. However, in some cases, institutional representatives were concerned about an inability to validate self-reported data. In addition, it may be difficult to be certain that past graduate students and postdoctoral researchers are correctly identified in Internet searches.

Some institutions reported having a database that allows faculty mentors and trainees to access their own information and update it as appropriate. However, compliance with requests to keep information up to date varies greatly. While such approaches may also be error prone, they are often less labor intensive for staff. Most institutions struggled with collecting career trajectory data, particularly past an initial position following graduation or completion of the training. This problem is particularly difficult for institutions relying on manual input of data.

**Postdoctoral researcher data are limited.**
Compared with PhD and MD/PhD program data fields, fewer postdoctoral fields are collected by the databases explored in this study, and even fewer data elements are made available to the public. Postdoctoral program career outcomes data are collected by fewer than half of the databases and for fewer years. While progress has been made on reforming the postdoctoral training system over the past decade, including the creation of postdoctoral training leadership positions or offices at institutions, some aspects, such as the collection of postdoctoral trainee data, have not significantly changed.[5]

**Databases are used in multiple ways.**
One of the primary institutional drivers for the creation of trainee databases is data collection for NIH training grant applications, and all 20 databases were used for this purpose. Many institutions pointed to the need to more broadly track the life cycle of a graduate student or postdoctoral researcher and keep track of trainee career outcomes to support program planning and evaluation, recruitment, and fundraising and development. Additional drivers for database development cited by institutions were the increase in the number of trainees entering programs who expect to have access to outcomes data and the potential for increased federal data-collection requirements.

*The research and research training communities must ensure that trainees are prepared to meet future biomedical workforce needs. Biomedical workforce data are vital for understanding the careers that trainees are entering, aligning training with those needs, and educating trainees about these career options. Institutions are encouraged to use this report as a resource to help facilitate local and national discussions around research trainee data collection.*

## ACKNOWLEDGMENTS

We thank the following institutions that participated in the study:

**Boston University School of Medicine**

**Brown University**

**Case Western Reserve University**

**Duke University School of Medicine**

**Emory University**

**Tufts University**

**University of Alabama at Birmingham**

**University of California, San Francisco**

**University of Texas Southwestern Medical Center**

**University of Virginia School of Medicine**

**Vanderbilt University Medical School**

**Yale University**

We thank the following members of the GREAT Group institutional approaches to tracking trainee information subgroup and AAMC staff for contributing to this project:

**Thomas Geoghegan, PhD, University of Louisville School of Medicine, Co-Chair**

**Naomi Rosenberg, PhD, Tufts University School of Medicine, Co-Chair**

**Myles Akabas, MD, PhD, Albert Einstein School of Medicine**

**John Alvaro, PhD, Yale University School of Medicine**

**Mary Bradley, MLA, Washington University in St. Louis, School of Medicine**

**Deirdre Brekken, PhD, University of Texas Southwestern Medical Center**

**John Horn, PhD, University of Pittsburgh School of Medicine**

**Ambika Mathur, PhD, Wayne State University**

**Dianna Milewicz, MD, PhD, University of Texas Health Science Center at Houston and M.D. Anderson Cancer Center**

**Molly Starback, MLS, Duke University School of Medicine**

**Jodi Yellin, PhD, AAMC**

**Anurupa Dev, PhD, AAMC**

RETURN TO RESULTS

**Figure 2. Data fields collected by surveyed institutions. The color gradient from green to red corresponds with numbers from high to low. Eleven institutions reported collecting PhD program data within at least one of their databases, 11 collected MD/PhD program data, and 10 collected postdoctoral program data. NRSA = Ruth L. Kirschstein National Research Service Award; TG = training grant; FTTP = full-time training position.**

## FIGURE 2A. Program characteristics.

| PhD students | Number of institutions collecting the data field |
|---|---|
| Number of programs | 11 |
| Total number of students enrolled | 11 |
| Number of women enrolled | 11 |
| Number enrolled by race/ethnicity | 11 |
| Number of students in NIH underrepresented population* | 11 |
| Number of U.S. citizens or permanent residents enrolled | 11 |
| Time to PhD degree for each graduate | 11 |
| Number of matriculants that graduate within 10 years | 11 |
| Number of students who left the program with no degree | 11 |
| Number of students who left the program with a master's degree | 11 |
| Number of degrees conferred in the previous academic year | 10 |
| Number of NRSA training grants supporting graduate programs | 8 |
| Number of FTTPs on TGs | 8 |
| Number of FTTPs supported by external competititve fellowships | 6 |
| Age of student at receipt of degree | 6 |

| MD/PhD students | Number of institutions collecting the data field |
|---|---|
| Total number of students enrolled | 11 |
| Number of women enrolled | 11 |
| Number enrolled by race/ethnicity | 11 |
| Number of students in NIH underrepresented population | 11 |
| Number of U.S. citizens or permanent residents enrolled | 11 |
| Time to dual degree for each graduate | 11 |
| Number of degrees conferred in the previous academic year | 11 |
| Number of students who left the program with no degree | 11 |
| Number of students who left the program with a master's degree | 11 |
| Number of matriculants that graduate within 10 years | 10 |
| Number of students who left the program with an MD degree only | 9 |
| Number of students who left the program with a PhD degree only | 9 |
| Number of FTTPs supported by external competitive fellowships | 8 |
| Number of FTTPs | 7 |
| Age of student at receipt of degree | 7 |

| Postdocs | Number of institutions collecting the data field |
|---|---|
| Total number of postdocs appointed | 9 |
| Number of female postdocs appointed | 9 |
| Number of postdocs appointed by race/ethnicity | 9 |
| Number of U.S. citizen or permanent resident postdocs | 9 |
| Length of training for each postdoc | 8 |
| Number of students in NIH underrepresented population | 7 |
| Number of NRSA training grants supporting postdocs | 6 |
| Number of FTTPs supported by external competitive fellowships | 6 |
| Number of FTTPs on TGs | 5 |

* Link to definition of NIH underrepresented population: http://grants.nih.gov/grants/guide/notice-files/NOT-OD-15-053.html.

RETURN TO RESULTS

## FIGURE 2B. Faculty characteristics.

| PhD students | Number of institutions collecting the data field |
|---|---|
| Total number of faculty appointed to programs | 8 |
| Number of female faculty appointed to programs | 7 |
| Number of faculty appointed to programs by race/ethnicity | 6 |
| Number of basic science faculty appointed to programs | 5 |
| Number of clinical science faculty appointed to programs | 5 |
| Estimated number of faculty eligible to be thesis advisors | 3 |

| MD/PhD students | Number of institutions collecting the data field |
|---|---|
| Total number of faculty appointed to program | 8 |
| Number of female faculty appointed to program | 7 |
| Number of faculty appointed to programs by race/ethnicity | 6 |
| Number of basic science faculty appointed to program | 5 |
| Number of clinical science faculty appointed to program | 5 |
| Estimated number of faculty eligible to be thesis advisors | 5 |

| Postdocs | Number of institutions collecting the data field |
|---|---|
| Number of faculty with postdocs | 6 |
| Number of female faculty with postdocs | 6 |
| Number of basic science faculty with postdocs | 5 |
| Number of clinical science faculty with postdocs | 4 |
| Number of faculty with postdocs by race/ethnicity | 3 |

## FIGURE 2C. Incoming trainee populations.

| PhD students | Number of institutions collecting the data field |
|---|---|
| Total number of matriculating students** | 11 |
| Number of matriculating female students** | 11 |
| Number of matriculating students by race/ethnicity** | 11 |
| Number of matriculating U.S. citizens and permanent residents** | 11 |
| Number of applicants offered admission who matriculated** | 11 |
| GRE scores for each matriculant | 11 |
| Matriculating students in NIH underrepresented populations | 10 |
| Number of completed applications received each year | 10 |
| GRE scores for each applicant | 10 |
| Undergraduate GPA in science-related courses | 7 |

| MD/PhD students | Number of institutions collecting the data field |
|---|---|
| Total number of matriculating students** | 11 |
| Number of matriculating female students** | 11 |
| Number of matriculating students by race/ethnicity** | 11 |
| Matriculating students in NIH underrepresented populations | 11 |
| Number of matriculating U.S. citizens and permanent residents | 11 |
| Number of applicants offered admission who matriculated** | 10 |
| MCAT scores for each matriculant | 10 |
| Number of applicants who received offers of admission** | 9 |
| Number of completed applications received each year | 8 |
| MCAT scores for each applicant | 8 |
| Undergraduate GPA in science-related courses | 7 |

| Postdocs | Number of institutions collecting the data field |
|---|---|
| Total number of first-year postdocs** | 9 |
| Number of incoming female postdocs** | 9 |
| Number of matriculating U.S. citizens and permanent residents | 9 |
| Number of incoming postdocs by race/ethnicty** | 8 |
| Incoming postdocs in NIH underrepresented populations | 7 |
| Age of incoming postdoctoral appointees | 7 |

** Data are collected each year.

RETURN TO RESULTS

## FIGURE 2D. Publications.

| PhD students | Number of institutions collecting the data field |
|---|---|
| Number of first-author publications for each graduate | 7 |
| Number of publications other than first-author per graduate | 7 |

| MD/PhD students | Number of institutions collecting the data field |
|---|---|
| Number of first-author publications for each graduate | 7 |
| Number of publications other than first-author per graduate | 7 |

| Postdocs | Number of institutions collecting the data field |
|---|---|
| Number of first-author publications for each postdoc | 2 |
| Number of publications other than first-author per postdoc | 2 |

## FIGURE 2E. Career outcomes.

| PhD students | Number of institutions collecting the data field |
|---|---|
| Number of all graduates in different career sectors | 7 |
| Number of graduates in the past 5 years in different career sectors | 7 |
| Number in tenured or tenure-track academic positions | 2 |
| Number of graduates in the past 5 years in tenure-track positions | 2 |

| MD/PhD students | Number of institutions collecting the data field |
|---|---|
| Number of all graduates in different career sectors | 7 |
| Number of graduates in the past 5 years in different career sectors | 7 |
| Number in tenured or tenure-track academic positions | 3 |
| Number of graduates in the past 5 years in tenure-track positions | 3 |

| Postdocs | Number of institutions collecting the data field |
|---|---|
| Number of all postdocs in different career sectors | 4 |
| Number of postdocs in the past 5 years in different career sectors | 4 |
| Number of tenured or tenure-track academic positions | 2 |
| Number of graduates in the past 5 years in tenure-track positions | 2 |

## FIGURE 2F. Public information.

| PhD students | Number of institutions collecting the data field |
|---|---|
| Average time to degree | 9 |
| Alumni career outcomes | 7 |
| Number of graduating students | 6 |
| Number of applications received | 5 |
| Attrition rate | 5 |
| Number of offers made | 4 |
| Number of entering students | 3 |
| Number of students leaving after a master's | 3 |
| Number of interviews granted | 2 |

| MD/PhD students | Number of institutions collecting the data field |
|---|---|
| Alumni career outcomes | 7 |
| Number of graduating students | 6 |
| Number of applications received | 5 |
| Number of interviews granted | 4 |
| Attrition rate | 4 |
| Average time to dual degree | 4 |
| Number of entering students | 3 |
| Number of offers made | 2 |
| Number of students leaving with a PhD only | 2 |
| Number of students leaving with an MD only | 2 |

| Postdocs | Number of institutions collecting the data field |
|---|---|
| Number of current postdocs | 2 |
| Average length of training | 1 |
| Alumni career outcomes | 1 |

## NOTES


1. National Institutes of Health. 2012. Biomedical Research Workforce Working Group Report. Bethesda, MD: National Institutes of Health.

2. National Institutes of Health. 2014. Physician-Scientist Workforce Working Group Report. Bethesda, MD: National Institutes of Health.

3. Alumn, J.R., Kent, J.D., and McCarthy, M.T. 2014. Understanding PhD Career Pathways for Program Improvement: A CGS Report. Washington, DC: Council of Graduate Schools.

4. The GREAT Group is a professional development group for graduate deans, MD/PhD program directors, and postdoctoral program directors who have the responsibility for PhD, MD/PhD, and postdoctoral training within medical schools and teaching hospitals.

5. National Academies. 2014. *The Postdoctoral Experience Revisited*. Washington, DC: National Academies Press.


**Association of**
**American Medical Colleges**
655 K Street, N.W., Suite 100, Washington, D.C. 20001-2399
**T** 202 828 0400
www.aamc.org