REPORT

Final Report

Evaluation of the NHLBI Programs of Excellence in Molecular Biology

To

National Institutes of Health

National Heart, Lung, and Blood Institute

Bethesda, Maryland 20892

December 1997

Battelle
. . . Putting Technology To Work

Final Report

Contract No. 292-92-0053

on

Evaluation of the NHLBI Programs of Excellence in Molecular Biology

to

Carl A. Roth, Ph.D., LL.M.
Associate Director for
Scientific Program Operation
National Heart, Lung, and Blood Institute
National Institutes of Health
Bethesda, Maryland 20892

December 1997

by

George Provenzano, Ph.D.
Project Director

Joanne Abed, Ph.D.
Robert Orwin, Ph.D.
Bennett Pierce, M.S.
John M. Rose, M.A.
Shawn Shumaker, M.A.S.
Valerie F. Williams, M.A., M.S.

BATTELLE
Centers for Public Health Research and Evaluation
2101 Wilson Boulevard, Suite 800
Arlington, VA 22201

Contents

Tables

Figures

Appendices

1.0 Introduction

1.0 Introduction

In this introductory chapter we present (1) background information on the Programs of Excellence in Molecular Biology, (2) the study purpose, (3) a summary description of project activities, and (4) a brief outline of the report.

1.1 Background

Over eight years ago, the National Heart, Lung, and Blood Institute (NHLBI) convened a committee of experts to examine the extent to which the techniques of modern molecular biology were being applied in areas related to heart, lung, and blood diseases. The committee concluded that applications of molecular biology in cardiovascular and pulmonary areas were lagging behind applications to other clinical problems. Based on this finding, the committee recommended that the NHLBI develop and implement a new grant program to support concerted research efforts to apply molecular biology techniques to fundamental cardiovascular, pulmonary, and related hematologic research.

The NHLBI responded by developing a Request for Applications (RFA) for the Programs of Excellence in Molecular Biology (POEMB). The objectives established for the POEMB and announced in the RFA were to:

- Stimulate the use of molecular biology approaches in areas consistent with the mission of the NHLBI where the new techniques of molecular biology have been underutilized and
- Provide opportunities for investigators who have the potential for establishing or redirecting research careers to become skilled in the experimental strategies and techniques of molecular biology and their application to research relevant to the mission of the NHLBI.

The POEMB grants were awarded as program project grants (P01s) consisting of multidisciplinary teams of independent investigators focused on a central unifying theme.

Five features were listed in the RFA as key aspects of the POEMB grant program: (1) the seven-year award, (2) support for recruiting essential scientific expertise, (3) support for institutional

environment and resources, (4) support for the use of experimental design and methods, and (5) support for new investigators. Each of these is described in further detail below.

- **The seven-year award.** Each POEMB was funded for an initial period of seven years but afforded an opportunity for the institutions who had received the awards to submit a special renewal application for one additional seven-year period. The extended award period and simplified renewal option were adopted to enable a POEMB to be innovative, to pursue new developments in the rapidly advancing field of molecular biology, and to embark on the application of molecular biology to complex experimental systems and new experimental models that would require additional time to develop and use.

- **Support for recruiting essential scientific expertise.** In order to ensure the most effective combination of scientific disciplines, POEMB applicants were allowed to include a request for funds to recruit faculty to augment or strengthen the skills, expertise, and capabilities of existing faculty.

- **Support for institutional environment and resources.** Applicants were also permitted to request funds for incidental alteration and renovation of facilities and for equipment needed to conduct research using the technologies of molecular biology.

- **Support for the use of experimental design and methods.** In an effort to provide broader degrees of research freedom, and to encourage innovative approaches, applicants were only required to provide brief descriptions of preferred and alternative experimental approaches, strategies, and proposed research directions.

- **Support for new investigators.** To provide for the development of researchers in areas related to cardiovascular and pulmonary diseases able to use the technologies of molecular biology effectively, applicants were allowed to request support for young investigators or investigators new to the discipline of molecular biology. Upon completion of their research duties, it was expected that the new investigators supported by POEMB would be able to establish independent laboratories devoted to the application of molecular biological strategies and approaches to fundamental cardiovascular, pulmonary, or related hematology research problems.

Because the use of molecular and cellular biology in the study of hematologic disease and related phenomena was both well-established and productive in 1986 when the initiative was being formulated, applications for support of programs that focused mainly on hematologic questions were not considered to be responsive to the RFA. However, collaborative approaches between hematologists and cardiovascular and pulmonary research investigators were permitted in areas of study of cardiovascular and pulmonary mechanisms that overlap significantly with hematologic interests as long as the major thrust remained focused on normal or altered cardiovascular or pulmonary function.

The NHLBI ultimately supported three POEMBs, two awarded in 1988 (to the Massachusetts Institute of Technology and the University of Cincinnati) and one awarded in 1989 (to the University of California at San Francisco). At the time of the release of this report, only MIT and University of Cincinnati are still receiving funding. Funding to the University of California at San Francisco ended in November 1996. Average annual support for each of the POEMBs has exceeded $2 million since fiscal year 1988. NHLBI support for the POEMBs totaled $6.7 million in fiscal year 1993.

1.2 Study Purpose

At this time, a comprehensive evaluation of POEMB is warranted. With POEMB having completed its first seven-year cycle, sufficient time has elapsed for objective evidence of the program's impact to become available. Moreover, the NHLBI is presently considering ways to stimulate research on multifactorial conditions such as atherosclerosis and hypertension. If the POEMB approach is shown to be effective, it will constitute a model for the NHLBI and for other Institutes, Centers, and Divisions (ICDs) at NIH interested in stimulating research activity and progress in these and other areas at the frontier of biomedical science.

Anecdotal evidence, available from progress reports and presentations made in sponsored sessions at national meetings, offers some evidence of the impact of the POEMB initiative. For example, the first reported *in vivo* application of antisense technology was performed in one of the POEMBs. In addition, special sessions have been held in conjunction with three major clinical research meetings (two Tri-Societies and one American Heart Association) to present results of POEMB research, all of which were extraordinarily well attended. A review of each of the individual POEMBs will be conducted as part of the renewal process, but those reviews will not provide an overview of the POEMB approach and, to date, there has been no systematic effort to document its effectiveness.

This evaluation focuses upon measures of the extent to which the stated goals of the initiative have been realized. Thus, we seek to characterize research activity in cardiovascular and pulmonary research using the techniques of molecular biology in terms of research grant activity and publications, both before and after implementation of the POEMB, produced by groups who are supported by the POEMB and those who are not. Qualitative data were also collected on the effectiveness of the POEMB approach and on research accomplishments by POEMB applicants.

1.3 Project Activities

Activities on this project included (1) preparation of data sets, (2) determining specialty areas of investigators and trainees, (3) determining research activities of investigators and trainees, (4) investigating the missionary impact on departing staff, and (5) identifying advantages/disadvantages and accomplishments of the POEMB concept.

1.3.1 Preparation of Data Sets

Battelle prepared two types of data sets for this project: (1) a person-based data set of all individuals cited in both successful and unsuccessful applications and progress reports and (2) analysis data sets.

1.3.1.1 Battelle Person-based Database

From POEMB applications and progress reports a data set was prepared containing all relevant, available identifiers (e.g., full name, SS#, DOB, doctoral level degrees, granting institution of doctoral degree) of proposed and (in the case of progress reports) actual program staff. Detailed information was considered especially critical for individuals with very common names and/or individuals with missing SS#s but was collected in the same thorough manner for all staff cited.

Because the data provided on any given individual in one report was often incomplete and/or inconsistent with the data provided in a subsequent report, resolving inconsistencies and ambiguities in the data was an important aspect of this task.

Although the original workplan called for the development of two distinct data sets (one for proposed staff cited in applications and one for trainees cited in progress reports), we elected to create one core data set that could be subdivided in several ways through sorting on a status code (SCODE) field. We had several reasons for doing this:

1. The fields required for the proposed staff and the trainee data sets were identical. We needed to collect the same information on all staff cited, regardless of their status.

2 Through sorting on the SCODE field, the two groups could be separated for analysis whenever desired. Use of a sort field also enabled us to differentiate more finely. Rather than the two groups proposed in the original workplan, we differentiated five possible statuses: PI (status 1), key investigators (status 2), potential trainees (status 3), trainees (status 4), and support staff (status 5).

3. Data entry for all statuses could be completed in one pass through the POEMB documents. Data cleaning was also simplified by having only one data set to edit.

4. Status distinctions among POEMB staff cited in applications and progress reports were not always readily apparent the first time an individual's name appeared. For example, trainees were sometimes cited without their academic degrees and so were initially categorized as status 5 (support staff). In several cases, we did not learn until verification of the data by the applicant institution that the individual had a recent MD or PhD and was therefore a trainee.

Chapter 2 of this report describes the procedures followed in preparing the person-based data set.

1.3.1.2 Creation of Analysis Data Sets

Analysis data sets were created by "matching" the information in the person-based data sets to records in the: Consolidated Grant Applicant File (**CGAF**), Trainee and Fellow File (**TFF**), Doctorate Records File (**DRF**), Computer Retrieval of Information on Scientific Projects (**CRISP**), American Heart Association Grant Files (**AHA**), and **Medline**. Due to the large number of records in these databases, the matching involved a combination of an automated algorithm with follow-up hand verification for a subgroup of records. The automated algorithm was based upon variations of a person's last name, first name, date of birth, and social security number. Hand verification was performed for records that only matched on variations of the person's last name and first name or that did not match on any criteria employed in the automated algorithm.

Once the matching had been performed, several of the resulting data sets were modified to facilitate the statistical analyses. In particular, variables were created to indicate which records corresponded to persons cited in successful POEMB applications. Additionally, variables were created to indicate the time frame, relative to the POEMB application, for grant applications (**CGAF**), journal articles (**Medline**), or **AHA** awards. Other variables were created for specific databases. A complete description of the creation of each analysis database is provided in Chapter 2.

1.3.2 Determine Specialty Area of Training of Investigators and Trainees and Previous NRSA Postdoctoral Training Received by POEMB Trainees

From the **DRF**, we identified the area of doctoral-level training for investigators cited on POEMB applications and trainees reported in POEMB progress reports. For those cases in which SS#s could not be found in the POEMB applications or progress reports, matches using the last name and the first initial of the first name were attempted. The full names of those matches were then examined manually to determine whether any were valid.

By tracking the extent of postdoctoral training received by POEMB trainees using the NIH **TFF**, the NIH IMPAC system, and the NIH **TAF** and by searching the **CGAF** for prior or concurrent receipt of NIH career development awards or research grants, we sought to determine:

- Whether POEMB training is instead of or in addition to postdoctoral NRSA training and/or NIH career development awards and
- Whether POEMB training has been provided to any individuals who have already achieved independent investigator status.

For those cases in which SS#s could not be found in the POEMB applications or progress reports, matches using the last name and first initial of the first name were attempted. Full names of those matches were then examined manually to determine whether any were valid.

An interim deliverable was produced (*Interim Report for Task 3: Training of Investigators and Trainees*) summarizing (1) the fields of training for individuals named as staff in POEMB applications and for individuals trained under the POEMB awards and (2) the extent to which individuals receiving training on POEMB have either received NRSA training support, an NIH career development award, or an NIH research grant. Activities and findings related to this task are summarized in Chapter 3 of this report.

1.3.3 Determine the Research Activity of Individuals Cited in Successful and Unsuccessful POEMB Applications and Progress Reports

NIH grant application and award sources were used to identify the nature of NIH grant activity (assigned ICDs and research areas) before and after the date of the POEMB application for the

individuals included on POEMB applications and for POEMB trainees after their training experience. Grant applications were classified as HL (heart/lung) or non-HL using the Institute/Center/Division (ICD codes).[1] Using the **CGAF** baseline data file for the years 1984 to 1993, the percentage of investigators that submitted one or more ICD 'HL' or 'non-HL' grant application was determined. This was done separately for "before" and "after" POEMB application date and for each time period. In order to determine the impact of POEMB on research direction/focus, the percentage of investigators who submitted one or more 'non-HL' applications "before" POEMB application date and one or more 'HL' applications "after" POEMB application date was calculated for both POEMB applicant groups. In addition, these data were modeled statistically, and the impact of POEMB on the dependent variables described above was assessed. **CRISP** descriptors[2] were used to classify grants as heart/lung related (HL), molecular biology related (MB), heart/lung and molecular biology related (HL/MB) or "other" (CC). In order to determine the impact of POEMB on research direction/focus, the percentage of investigators in both applicant groups who were awarded one or more grants characterized in the following categories were calculated: 'HL' "before" POEMB application and one or more grants characterized as 'MB' "after" POEMB application; 'HL' "before" and 'HL/MB' "after"; 'MB' "before" and 'HL' "after"; 'MB' "before" and 'HL/MB' "after"; and 'other' "before" and 'HL/MB' "after". These data were modeled statistically, and the impact of POEMB on the dependent variables described above was assessed.

Using **AHA** (American Heart Association) award data, the AHA activity was determined for the POEMB applicants before and after the POEMB application and for the POEMB trainees subsequent to their training by calculating the percentage of investigators who were awarded one or more AHA awards. This was done separately for "before" and "after" POEMB application date and for each time period. In order to determine the impact of POEMB on research direction/focus, the percentage of investigators who received no AHA awards "before" the POEMB application date and one or more awards "after" POEMB application date was calculated separately for both POEMB applicant groups. In addition, the extent of overlap in research personnel and trainees between the AHA-Bugher Awards

---

1 **ICD** codes indicate the institution to which the application was submitted. "HL" indicates grant applications submitted to NHLBI or one of its divisions.

2 **CRISP** descriptors are subject terms that most appropriately describe each research project in terms of the focus and techniques used.

and the POEMBs was determined. The percent of overlap with POEMB awardees was then compared with the corresponding percentage for the unsuccessful POEMB applicants. The timing of the overlap was noted in terms of whether the overlap occurred prior to, concurrently with, or subsequent to the POEMB application. The results of the activities described above are discussed in Chapter 4.

Chapter 5 reports on activities that involved using the original POEMB grant applications, the subsequent progress reports for POEMB awardee groups, and **Medline** to determine the research journal publication activity for the individuals included on POEMB applications before and after the POEMB application and for the POEMB trainees subsequent to their training. The percentage of investigators that published one or more articles was calculated for researchers in both POEMB applicant groups for "before" and "after" POEMB application date and for each time period. Journal articles were classified into the following five categories based upon journal title: cardiovascular diseases, respiratory diseases, molecular/microbiology, and "other". The percentage of investigators that published articles in each of these categories was calculated separately for each POEMB applicant group and for each time period. **Medline** MeSH subject terms were then used to classify each article as strictly heart/lung related (HL), strictly molecular biology related (MB), strictly "other", or a combination of heart/lung and molecular biology related (HL/MB). The following percentages were calculated for investigators in both POEMB applicant groups that published (1) one or more articles with HL-designated MeSH terms; (2) one or more articles with MB-designated MeSH terms; and (3) one or more articles with HL/MB-designated MeSH terms. In order to determine the impact of POEMB on the direction/focus of the applicants' research, we then examined the degree to which interactions exist between journal title and MeSH terms (HL/MB overlap). We calculated the number of investigators in both POEMB applicant groups that published articles in: (1) HL-designated journals with MB MeSH terms; (2) MB-designated journals with HL MeSH terms; and (3) in "other" journals with HL/MB-designated MeSH terms. These data were also modeled statistically, and the impact of POEMB on the dependent variables described above was assessed.

1.3.4 Investigate the Missionary Impact of Departed POEMB Staff and POEMB Trainees

Using these same data (**CGAF**, **CRISP**, and **Medline**), the grant and publication activity of investigators and trainees who left the POEMB was examined in an effort to determine whether these

individuals continued to pursue applications of molecular biology in the fields of cardiovascular and/or pulmonary disease. It was assumed that the subsequent grant applications and publications of former POEMB staff and POEMB trainees would show whether these researchers retained an interest in molecular biology applications to cardiovascular and/or pulmonary disease after leaving POEMB. These analyses are discussed in Chapters 4 and 5 of this report.

1.3.5 Identify Advantages and Disadvantages and Outstanding Accomplishments of the POEMB Concept

Battelle developed separate guides for interviews conducted with each of the three following groups: (1) principal investigators on successful POEMB applications, (2) principal investigators on unsuccessful POEMB applications, and (3) other investigators on successful POEMB applications. Interviews were then conducted with each of the three POEMB principal investigators, with nine other POEMB investigators, and with nine of the PIs on unsuccessful applications. We sought to obtain the interviewees' perspective as to the value and desirability and the advantages and disadvantages of the POEMB approach, with particular emphasis on the aspects of the program that led them to organize an application or to participate in its preparation.

An interim report was delivered entitled *Identifying the Advantages, Disadvantages, and Accomplishments of the POEMB*. Activities and findings from this phase of the project are discussed in Chapter 6 of this report.

1.4 Organization of the Report

This report is divided into six chapters. In this chapter, we have presented background information on the POEMB program and a summary of the tasks and activities undertaken in connection with this evaluation. Chapter 2 describes the data set developed by Battelle and used as a basis for the record matching done to support this evaluation. In Chapter 3, we provide descriptive statistics on the research institutions that responded to NHLBI's Request for Applications (RFA) for the POEMB and describe the POEMB trainees in terms of their specialty areas and their previous post-doctoral training.

The report continues by assessing how effective POEMB has been in stimulating new molecular research in cardiovascular and pulmonary disease. In Chapter 4, we compare grant application and award activity between successful and unsuccessful POEMB applicants. In Chapter 5, we compare successful and unsuccessful POEMB applicants in terms of their research journal publication efforts.

Lastly, Chapter 6 presents qualitative findings from interviews with POEMB principal investigators, senior investigators, and their counterparts on unsuccessful applications. The findings highlight the advantages and disadvantages of the POEMB approach to stimulating research activity and progress achieved in this area of biomedical science by POEMB grantees.

2.0 Preparation of Data Sets

2.0 Preparation of Data Sets

In order to assess the impact of the POEMB grant program on the careers of those individuals associated with it, we first needed to compile a data set of all individuals cited in POEMB applications and progress reports. This person-based data set had to include as many identifying data as possible (name, Social Security number, date of birth, earned degrees) to aid in the subsequent verification of matches with larger data sets. We then needed to match the person-based data set against education, grant, and publication data sets to create analysis data sets consisting of subsets of the larger data sets that pertained only to individuals associated with POEMB (either as grantees or as unsuccessful applicants).

In this chapter, we describe the extensive preliminary work required for this study, which involved the preparation of two different types of data sets: (1) a person-based data set of all individuals cited in POEMB applications and progress reports and (2) analysis data sets derived from the matching of the person-based data set against larger NIH and other related data sets of interest to the study. We address each of these preparatory activities in turn.

2.1 Preparation of the POEMB Person-Based Data Set

Our goal in designing the POEMB person-based data set was to cull and store electronically as much relevant information as possible from the POEMB applications and progress reports, which represented our only documentary data sources for the names and other identifying information on individuals associated with the POEMB grant program. Below we describe (1) data sources and data entry procedures, (2) data verification and cleaning procedures, and (3) the structure of the resulting data set.

2.1.1 Data Sources and Data Entry Procedures

During initial data entry, we first reviewed applications and then grantee progress reports. We created a new record in **PMBPERS**[1] each time a new name appeared in a given application, assigning the record a unique number (or PCODE) and entering any identifying information (e.g., date of birth, Social Security number) that accompanied the first citation. If a name appeared more than once in the same application or (in the case of successful grantees) appeared again in a subsequent progress report, information accompanying each subsequent citation was checked against that already in the person record, any new information was added, and any discrepancies were noted on the original record. If a name appeared in a progress report that had not been present in the original application, a new record was added to **PMBPERS**.

Initially to ensure that all persons were included in the database, we allowed a name to be entered more than once in **PMBPERS** in connection with multiple applications (for example, when an institution submitted both a Round 1 and a Round 2 application citing some of the same personnel). In this case, the two records were assigned different PCODE's. Such duplicates were later culled from the data file, through a procedure described in Section 2.1.2.2 below.

The two types of documentary data sources available to Battelle for constructing the person-based data set were (1) POEMB applications and (2) POEMB progress reports. Applications were available for three successful and 21 unsuccessful applicant institutions; progress reports were available for successful applicant institutions only. Below we describe our data entry procedures for each type of data source.

2.1.1.1 POEMB Applications

Of the 27 POEMB applications received by NHLBI in response to the POEMB RFA, Battelle received a total of 24 applications from NHLBI for use in this study. The remaining three applications could not be located in the NHLBI archives.[2] Table 2.1 summarizes the POEMB applications reviewed by

---

1 Throughout this chapter we will be denoting the person-based data set using the acronym **PMBPERS**.

2 Battelle did not receive applications for Rottman (41486), Shafer (41481), or Maciag (41505).

Table 2.1
POEMB Applications Reviewed for This Evaluation

App#	NHL#	Date	Institution	Dept	PI
01	41484	1988	MIT	Biology	Rosenberg
02	41496	1988	U Cincinnati	Microbiology	Lingrel
03	43821	1989	UCSF	Cardiovascular Research Institute	Bourne
04	41485	1988	Washington U	Pediatrics	Colten
05	41503	1988	U Penn	Cardiovascular Pulmonary Div	Fishman
06**	43826	1989	UC San Diego	Medicine	Rosenfeld
07	43741	1989	U Chicago	Medicine	Page
08	43749	1989	U Illinois	Internal Medicine	Swartz
09	41489	1988	U Alabama/Birm	Medicine	Segrest
10	43740	1989	Washington U	Pediatrics	Colten
11	43739	1989	Baylor	Cell Biology	Means
12	43862	1989	Children's Hospital	Cardiology	Nadal-Ginard
13	41509	1988	Baylor	College of Medicine	Gotto
14*	43850	1989	U Washington	Pathology	Ross
15	41518	1988	Johns Hopkins	Biochemistry	Grossman
16	41519	1988	W Alton Jones Cell Science Center	Cell Science	Harris
17	41506	1988	U Illinois	Food Chemistry	Kummerow
18	41511	1988	UC San Diego	Pharmacology	Karin
19*	43580	1989	U Washington	Pathology	Ross
20**	43826	1989	UC San Diego	Medicine	Rosenfeld
21	41493	1988	Duke	Medical Center	Greenfield
22	41483	1988	U Iowa	Cardiovascular Center	Abboud
23	41504	1988	UCSF	Metabolic Research Unit	Baxter
24	41494	1988	Children's Hospital	Cardiology	Nadal-Ginard
25	41497	1988	SUNY/Syracuse	Anatomy and Cell Biology	Foster
26	41482	1988	Boston U	School of Medicine	Brody

* Applications 14 and 19 were duplicates.
** Applications 6 and 20 were duplicates.

Battelle for this study. Below we describe the components of the applications from which we were able to draw names and important identifying information on proposed and actual POEMB staff. These include: (1) cover page, (2) senior investigators table, (3) biosketches, and (4) proposed research component descriptions.

Cover page. The cover page of each application contained the following information, which was entered into a summary file called **PMBAPP**:

- Institution applying for the POEMB grant
- NIH-assigned application number
- Social Security number of Principal Investigator (PI)
- Date application was received by NHLBI
- Review score assigned by Grant Review Panel

The cover page of each application contained the following information, which was entered into a summary file called **PMBPERS**:

- Full name (including initials) of PI
- Title of PI
- Academic degrees of PI
- Social Security number of PI

The PI was assigned Status 1 (Principal Investigator) in a field called SCODE.

The cover page of each application contained the following information, which was entered into a summary file called **PMBSITE**:

- Mailing address for each "performance site" proposed
- Telephone number for principal performance site

Senior Investigators Table. A required element of the POEMB application was the Senior Investigators Table. This section of the application contained the following data elements for each individual being proposed as a senior investigator, which were entered into the **PMBPERS** table:

- Full name (including initials)
- Advanced degrees
- Social Security number
- Academic title and/or proposed role on the grant

With the exception of the Principal Investigator who was assigned Status 1, all Senior Investigators were assigned Status 2 (Senior Investigator) in the field SCODE, unless an advanced degree was not provided, in which case they were assigned Status 5 (Support Staff).

Biosketches. A second required element in the POEMB application was the Biosketch Section, which provided biosketches on senior investigators being proposed. Data elements from the biosketches that were entered into the **PMBPERS** file included:

- Full name (including initials)
- Date of birth (DOB)
- Position or title

Data elements from the biosketches that were entered into a summary file called **PMBED** included:

- Degree earned
- Degree field
- Institution from which degree was obtained
- Year conferred

Educational background data for each degree earned were entered as an individual record in **PMBED**, linked to a person-record in the **PMBPERS** table by means of a unique identifier, the PCODE (or person code). Any discrepancies noted between data as provided in the Senior Investigators Table and in the biosketches were noted in a DISCREPANCIES field in **PMBPERS**. The publications listing in the biosketch was scanned to determine the initials as commonly used in bibliographic citations, and all variant forms of citation found in the publications listing were noted in the DISCREPANCIES field.

Biosketches were often provided in the POEMB applications for staff members not listed in the Senior Investigators Table. If these individuals had advanced degrees (beyond the master's level), they were assigned Status 3 (Potential Trainee). If they possessed degrees of master's level or below, they were assigned Status 5 (Support Staff). If a biosketch appeared for a Status 3 individual in one of the three awarded applications and the individual was subsequently referred to as a junior investigator (or similar title) in a progress report, the individual's status was revised to Status 4 (Trainee).

Research component descriptions. POEMB applications routinely included Research Component Descriptions, which summarized activities to be conducted on each of the proposed research projects. These often included listings of proposed staff. New names were added to the **PMBPERS** table along with any accompanying information (degrees, position title, initials used in publication). Those with degrees of master's level and below or with no degree were assigned Status 5 (Support Staff), those with degrees above the master's level were assigned Status 3 (Potential Trainee), and those Status 3 individuals who subsequently appeared in a progress report by a successful applicant institution were later reassigned to Status 4 (Trainee). See Table 2.2 for a summary of the status codes (or SCODES) and their meanings.

Table 2.2 Description of Status Codes Assigned to Each Applicant

Status Code	Description	Reason for Classification
1	Principal Investigator	Individual listed as the Principal Investigator on the cover page of the application.
2	Senior Investigator	Individual listed in the Senior Investigators Table of the application, generally a PI of a proposed research component or otherwise an important investigator on the grant.[1]
3	Potential Trainee	Individual listed in a proposed research component of an application or in a biosketch with a degree beyond the master's level but not appearing in a subsequent progress report.
4	Trainee	Individual listed in a proposed research component of an application or in a biosketch with a degree beyond the master's level and appearing in a subsequent progress report.
5	Support Staff	Individual cited in an application or progress report with a degree of master's level or below or for whom no advanced degree is listed.[2]

---

1 In some applications *all individuals proposed* appeared in the Senior Investigators Table, even those without advanced degrees. Those with no advanced degrees were assigned Status 5 (Support Staff), despite the fact that they were listed in the Senior Investigators Table. All others were assigned Status 2 (Senior Investigators). Although it is conceivable that some recent MDs and PhDs might have been among them, we had no way of distinguishing between such Potential Trainees and "true" Senior Investigators and so were obliged to accept the institution's designation of them all as Senior Investigators.

2 *Lack of evidence* of an advanced degree was equated with lack of that degree, which means that some Status 3 or 4 individuals may have been misclassified as Status 5. Additional earned degree information received, either in a progress report or through institutional verification, occasionally allowed us to reclassify misclassified individuals. However, because of the possibility of misclassification, we included Status 5 individuals in our matching procedures.

2.1.1.2 Progress Reports

The three successful POEMB grantees were Rosenberg/Massachusetts Institute of Technology (41484), Lingrel/University of Cincinnati (41496), and Williams/University of California San Francisco (43821). Progress reports submitted by successful POEMB grantees over the years of their grant contained the following data sources of interest to this study: (1) biosketches, (2) budget sheets, (3) personnel listings, and (4) various supporting documentation, each representing additional sources of names and other identifying information.

Progress report biosketches. Biosketches appearing in progress reports were used as were biosketches accompanying applications to obtain data on additional staff who actually worked on research components of a POEMB grant. Any discrepancies between the data in these biosketches and those in previously reviewed data sources were noted in the DISCREPANCIES field of **PMBPERS**.

Budget sheets. The budget sheets submitted with the progress reports were used as a means to ascertain an individual's *actual* association with a POEMB grant (rather than merely their *proposed* association with a grant).

Fields were included in the **PMBPERS** table to enable the tracking of that association throughout the years of the grant. These fields were designated by 01, 02, ... 07 to denote the application (01) and up to six subsequent progress reports. An "X" in the field signified that the individual's name had appeared in the corresponding application or progress report. Any discrepancies in names or other identifying information were noted in the DISCREPANCIES field of **PMBPERS** along with the data sources among which the discrepancies occurred.

Personnel listings. Additional names were sometimes gleaned from various personnel listings in the progress reports. Often these listings were associated with a given research component. In later progress reports, more general staff listings were included that provided fairly complete data on personnel associated with the grant (including date of birth and Social Security number). Any discrepancies between these data and data from previously reviewed sources were noted in the DISCREPANCIES field of **PMBPERS** along with the data sources among which the discrepancies occurred.

Supporting documentation. Occasionally anecdotal information demonstrating the impact of the POEMB program on a given individual associated with it appeared among the supporting documentation accompanying the progress reports. For example, a letter was included with one progress report stating that a researcher was leaving the grantee institution to pursue his/her career at another institution. This type of anecdotal information was noted in the ANECDOTES field of **PMBPERS**.

2.1.2 Data Cleaning and Verification

Our data cleaning and verification procedures included (1) verification of data entry and edits, (2) identification of duplicate person-records, and (3) verification by applicant institutions. Each of these procedures is discussed in turn below.

2.1.2.1 Verification of Data Entry and Edits

After initial data entry was completed, all data were verified through a process of "double-reading," whereby one researcher read aloud from the original application or progress report and a second researcher proofed what was being read against printouts of the data as entered. This activity presented an opportunity to discuss data entry decisions, standardize entries, and note additional discrepancies among data sources.

Following data cleaning, we entered all discrepancies noted in the DISCREPANCIES field into a separate database file called **PMBALT**, which contains fields for alternative name spellings, initials, Social Security numbers, and dates of birth. **PMBALT** is linked to the original person-record in **PMBPERS** by a unique identifier, the PCODE.

A second round of verification was then undertaken to ensure that all adjustments made in the first round of verification had been properly entered.

2.1.2.2 Identifying Duplicate Person-records

Once initial data entry and editing was complete, a printout (alphabetical by last name) of the entire **PMBPERS** file was run, in order to detect the presence of any duplicate person-records. The following types of duplicate entries were identified:

A person was entered twice from within the same application. Two cases were identified, one in Application 4 and one in Application 16. In order to eliminate duplicate person-records resulting from the same individual having been entered twice in association with the same application, information from the duplicate entries was combined into the first entry, and the second entry was deleted. Because this type of duplication resulted from the fact that a person's name was mentioned more than once within a given application (indicating two distinct data sources, such as a Senior Investigators Table and a Biosketch) discrepancies did occur and were noted in the DISCREPANCIES field of **PMBPERS** and later entered into the **PMBALT** table as alternative fields against which to match.

The same application was entered twice. Applications 6 and 20 from the University of California/San Diego and Applications 14 and 19 from the University of Washington were duplicate applications (1 P01 HL43826-01 and 1 P01 HL4385-01, respectively). In order to eliminate duplicate person-records resulting from the same application having been entered twice, data from the duplicate entries were combined and entered into the earliest of the two records (i.e., the record with the lower ACODE [or Battelle-assigned application number]). When the person-based data set was prepared for matching with the larger data sets, only data from Applications 6 and 14 were used. Data from Applications 20 and 19 were deleted from the data set used for matching. No discrepancies resulted from this type of duplication as the two applications were identical (i.e., a single data source).[1]

[1] Application 20 was only partial (missing all pages from the Research Component Descriptions). The list of names associated with this application was therefore considerably shorter than that for its counterpart, Application 6, although both are essentially the same application.

A single institution submitted two separate applications to the POEMB program, one for Round 1 and one for Round 2, that is, one set of applications for 1988 and one for 1989 and named some staff members in both applications. We received two applications each for the University of Illinois (Applications 17 and 8), the University of California/San Diego (Applications 18 and 6/20), Baylor College of Medicine (Applications 13 and 11), and Washington University (Applications 4 and 10).

In order to eliminate duplicate person-records resulting from the same individual being cited in both Round 1 and Round 2 applications, data from the two entries were combined and entered into the record with the lower Battelle-assigned ACODE, or application number. When the person-based data set was prepared for matching with the larger data sets, only data from Applications 8, 5, 11, and 4 were used. Data from Applications 17, 18, 20, 13, and 10 were deleted from the data set used for matching. Because the two applications in which the duplicates occurred were different POEMB applications (i.e., two distinct data sources), discrepancies between the two did occur and were treated in the same way as discrepancies between a successful application and its subsequent progress reports. That is, they were noted in the DISCREPANCIES field in **PMBPERS** and later entered into the **PMBALT** table as alternative fields against which to match.

2.1.2.3 Verification by Institutions

Following completion of our own data cleaning and verification procedures, we mailed out summary reports to all principal investigators from applicant institutions containing the names and other identifying information we had culled from their POEMB applications and/or progress reports. A printout of the Battelle-developed person-based database was designed that enabled us to present each applicant institution with a customized listing of the staff members they had cited. For verification purposes, we sent these printouts to all 24 POEMB applicants whose materials were reviewed for this study. The verification packet included (1) a letter of introduction from Dr. Sonia Skarlatos of NHLBI explaining the study and requesting the principal investigator's cooperation, (2) a letter from the Battelle Project Leader explaining the nature of the assistance required, and (3) the verification forms as customized for each POEMB applicant.

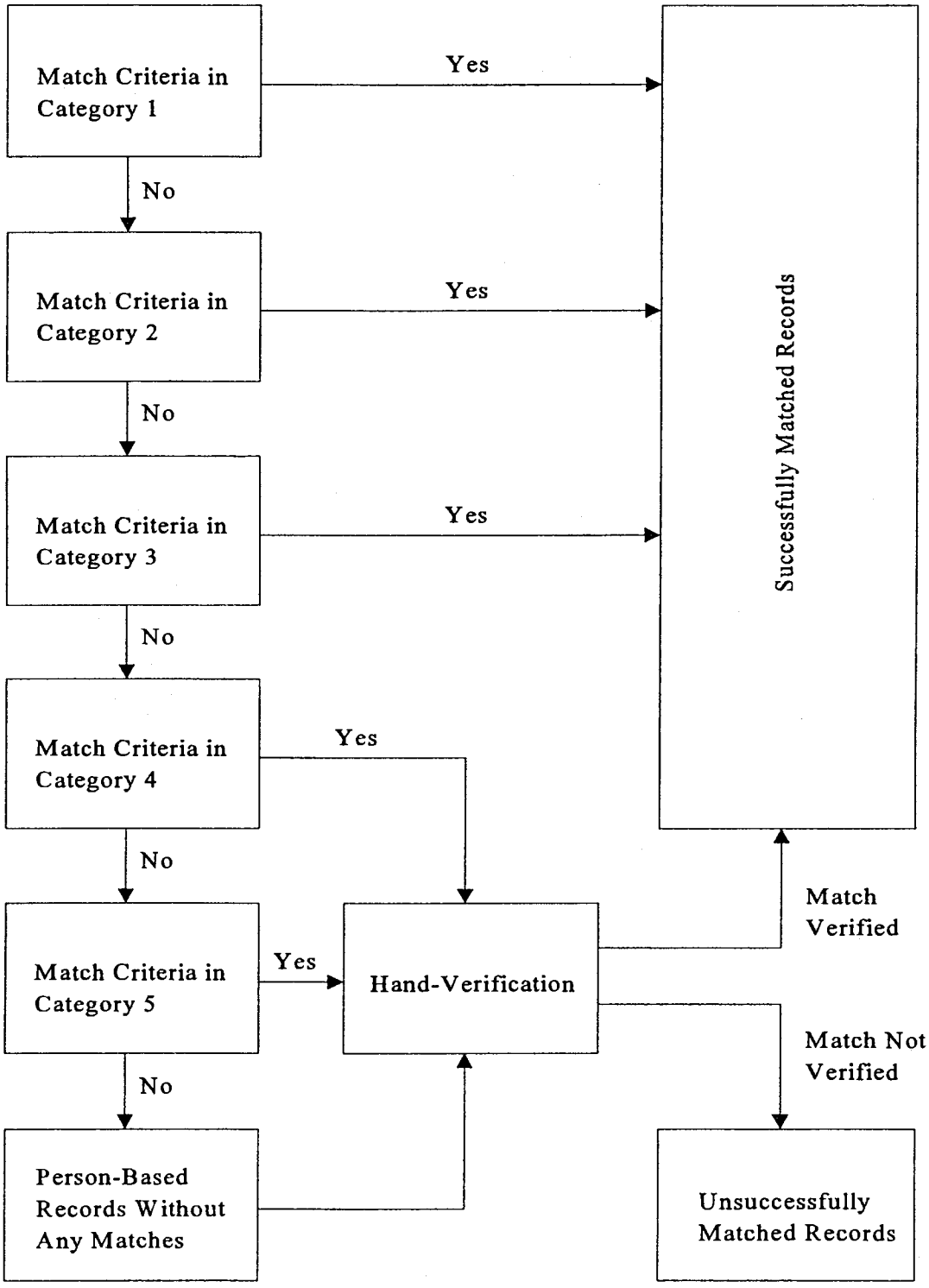
On the verification forms, problematic entries (i.e., those containing observed discrepancies) were separated from (1) those that simply had missing SS# and DOB and (2) those with complete information and for which no discrepancies had been identified. A "Discrepancy" column alerted institutions to any anomalies Battelle had identified between data from different data sources within the applications and/or progress reports. Blank columns alerted institutions to missing data elements (such as DOB or SS#). A "Comments" column was provided for institutions to note their revisions of the list and an OK column for them to check whether or not they had been able to verify a given entry as correct.

Data printouts were sent for institutional review in two separate batches. Phase 1 letters for Applications 1 through 20 were mailed out on March 30, 1995, with a response requested by April 10. Phase 2 letters for Applications 21 through 26 were mailed out on April 30, 1995, with a response requested by May 10, 1995.

Of the 24 verification lists sent out, 18 were returned. This was a much higher response rate than expected, and many of the institutions took great care to verify and amend the data. As responses were received, thank-you letters were mailed out to each institution that responded (from May 4 through June 20, 1995), with a copy to Dr. Carl Roth of NHLBI, the Delivery Order Officer on this contract.

Institutional additions and corrections were entered into a database similar in structure to **PMBALT** called **PMBINALT**. A printout of the newly received data was generated and used for the manual checking of questionable electronic matches. The supplemental data in **PMBINALT.DB** was also used in a second run of the "no-hit" records remaining unmatched after the first run against the NIH **INDEX** file, as will be described below.

Figure 2.1 Overall Methodology for Matching



activities are presented in Chapters 3, 4, and 5). These analysis data sets were created by "matching"[1] the information in the person-based data sets to records in the following data files:

- Consolidated Grant Applicant File (**CGAF**)
- Trainee and Fellow File (**TFF**)
- Doctorate Records File (**DRF**)
- Computer Retrieval of Information on Scientific Projects (**CRISP**)
- American Heart Association Grant Files (**AHA**)
- **Medline**

Section 2.2.1 describes the methodology that was used to perform the matching, while Section 2.2.2 provides additional detail regarding the creation of each analysis data set.

2.2.1 Methodology for Matching

The same general methodology was used to link people in the person-based data sets (**PMBPERS**) to their corresponding records in the **CGAF**, **TFF**, **DRF**, **CRISP**, **AHA**, and **Medline**. However, due to the nature of some of the data sets, the overall matching strategy was occasionally modified depending upon the data set in question. Section 2.2.1.1 presents an overview of the general matching strategy that was employed. Section 2.2.1.2 provides additional details concerning the matching to the data sets maintained by NIH (**CGAF**, **TFF**, **CRISP**) and the National Research Council (**DRF**), while Sections 2.2.1.3 and 2.2.1.4 provide additional details for matching to the **AHA** and **Medline** databases.

2.2.1.1 Overall Matching Methodology

The overall matching methodology consisted of two components: a computer algorithm and hand-verification. The computer algorithm was employed because of the extremely large number of records in the six data sets (e.g., the **CGAF** as maintained by NIH contained 1,550,206 records). Hand-verification was employed in cases where the computerized matching was questionable and as a quality assurance mechanism to spot-check the computer algorithm.

---

[1] Throughout the remainder of this report, the term "matching" refers to the practice of obtaining the appropriate information for the persons identified in the POEMB grant applications from other sources.

Initially, the computerized algorithm was based upon combinations of the following four primary fields in the person-based data sets:

- Last Name (LN);
- First Name (FN);
- Date of Birth (DOB); and
- Social Security Number (SSN).

However, the results of a pilot test indicated that additional fields needed to be developed to incorporate common misspellings and name changes or variations (e.g., Jon Doe versus Jonathan Doe) as well as discrepancies occurring between multiple data sources. Four additional fields were developed[1]:

- Alternative Last Name (LN2);
- Alternative First Name (FN2);
- Alternative Date of Birth (DOB2); and
- Alternative Social Security Number (SSN2).

These alternative fields did not exist for every record; they only existed if an alternative spelling or numeration was identified during the entry of the grant applicants (as described in Section 2.1).

Table 2.4 presents the matching criteria that were developed based upon combinations of the eight matching fields. Generally, the criteria can be separated into five hierarchical categories based upon the reliability of the matching. For example, the criteria in the first category require an exact match of four fields while the criteria in the second category require an exact match of only three fields.

Figure 2.1 illustrates the overall procedure used to match people in the person-based data sets to information in the other databases. As illustrated in the figure, records were sequentially matched to the five categories presented in Table 2.4. Records that successfully met one of the criteria in the first three categories were considered to be "true" matches and were not hand-verified. Records that matched at least one criterion in the last two categories were considered to be "possible" matches and were hand-checked to ensure that an appropriate match had been made. Finally, records in the person-based file that did not match to any of the criteria in Table 2.4 were also hand-checked.

---

[1] While there were other fields available, (middle name, degree, and application number) these fields were limited in their usefulness and were not used because they were either inconsistent or frequently missing.

Table 2.4 Computerized Matching Criteria

Category	Criteria
Category 1	1) LN, FN, DOB, SSN 2) LN, FN, DOB, SSN2 3) LN, FN, DOB2, SSN 4) LN, FN2, DOB, SSN 5) LN2, FN, DOB, SSN 6) LN, FN, DOB2, SSN2 7) LN, FN2, DOB, SSN2 8) LN2, FN, DOB, SSN2 9) LN, FN2, DOB2, SSN 10) LN2, FN, DOB2, SSN 11) LN2, FN2, DOB, SSN 12) LN, FN2, DOB2, SSN2 13) LN2, FN, DOB2, SSN2 14) LN2, FN2, DOB2, SSN 15) LN2, FN2, DOB, SSN2 16) LN2, FN2, DOB2, SSN2
Category 2	1) Same as Category 1 criteria 1-16, only using first initial (FI or FI2) for FN or FN2. 2) Same as Category 1 criteria 1-16, only using last initial (LI) for LN or LN2. 3) LN, FN, SSN 4) LN, FN, SSN2 5) LN, FN2, SSN 6) LN2, FN, SSN 7) LN, FN2, SSN2 8) LN2, FN, SSN2 9) LN2, FN2, SSN 10) LN2, FN2, SSN2
Category 3	1) Same as Category 2 criteria 3-10 criteria only using first initial (FI) for FN or FN2. 2) Same as Category 2 criteria 3-10 criteria only using last initial (LI) for LN or LN2. 3) LN, FN, DOB 4) LN, FN, DOB2 5) LN, FN2, DOB 6) LN2, FN, DOB 7) LN, FN2, DOB2 8) LN2, FN, DOB2 9) LN2, FN2, DOB 10) LN2, FN2, DOB2
Category 4	1) LN, FN 2) LN, FN2 3) LN2, FN 4) LN2, FN2
Category 5	1) LN, FI 2) LN, FI2 3) LN2, FI 4) LN2, FI2

2.2.1.2 Matching to NIH Databases

The **CGAF**, **TFF**, and **CRISP** were simultaneously matched through use of NIH's **INDEX** file. The **INDEX** file was developed, and is maintained, by NIH and was used in the creation and updates to the **CGAF** and the **TFF**. Furthermore, the **INDEX** file contains a unique *set number* that permits the linking of records from the **INDEX** file to every record in both the **TFF** and the **CGAF**. Therefore, the methodology outlined in Section 2.2.1.1 was employed to match the person-based records to the **INDEX** file and then the unique *set number* for each person was used to obtain the corresponding records in the **CGAF** and the **TFF**.

The appropriate records from the **CRISP** database were also obtained using the **INDEX** file. First, the records from the **CGAF** were extracted using the unique *set number*. Then, the *unique grant numbers* from each **CGAF** record were used to extract records from the CRISP database.[1]

Although the INDEX file was not used in the creation of the **DRF**, it contains a unique **DRF** identifier that can be used to link records in the **DRF** to those in the **INDEX** file. However, there are records in the **DRF** that do not have a corresponding record in the **INDEX** file. Thus, additional matching, using the methodology outlined in Section 2.2.1.1, was performed on the **DRF**.

Table 2.5 presents the results of matching to the **INDEX** file. The last three columns present the results of the matching in terms of the percentage matched in each matching category and the percentage of the total matched in each category. Table 2.6 presents the results of matching to the **DRF** file.

As might be expected, the results of the matching to the **INDEX** file were related to the indicated role of the person on the grant application. For example, a high percentage of principal investigators (SCODE 1) and senior investigators (SCODE 2) were matched to records in the **INDEX** database (100% and 96% respectively). However, this percentage was much lower (63%) for Trainees (SCODE 3) and even lower (48%) for Potential Trainees (SCODE 4). The results for SCODE 5 (Support Staff) were extremely low, and further matching for people in this category was not conducted.

---

[1] The **CRISP** data set contains information from awarded grants, while the **CGAF** contains information on all grant applications. Thus, every record in **CRISP** must also have a record in **CGAF**.

Table 2.5 Results of Matching to the NIH INDEX Database

Status	Matching Category	Number of Matches Found (% of Total Possible)		
		Successful Applications	Unsuccessful Applications	Total
Principal Investigator (SCODE 1)	1	75% (3/4)	87% (14/16)	85% (17/20)
	2	0% (0/4)	6% (1/16)	5% (1/20)
	3	0% (0/4)	6% (1/16)	5% (1/20)
	4	25% (1/4)	0% (0/16)	5% (1/20)
	5	0% (0/4)	0% (0/16)	0% (0/20)
	Total	**100% (4/4)**	**100% (16/16)**	**100% (20/20)**
Senior Investigator (SCODE 2)	1	77% (24/31)	72% (267/366)	73% (291/397)
	2	3% (1/31)	4% (18/366)	5% (19/397)
	3	6% (2/31)	18% (66/366)	17% (68/397)
	4	0% (0/31)	1% (3/366)	0% (3/397)
	5	0% (0/31)	0% (0/366)	0% (0/397)
	Total	**87% (27/31)**	**96% (354/366)**	**96% (381/397)**
Trainee (SCODE 3)	1	0% (0/32)	3% (8/220)	3% (8/252)
	2	0% (0/32)	0% (2/220)	0% (2/252)
	3	3% (1/32)	37% (82/220)	32% (83/252)
	4	43% (14/32)	15% (35/220)	19% (49/252)
	5	3% (1/32)	7% (16/220)	6% (17/252)
	Total	**50% (16/32)**	**65% (143/220)**	**63% (159/252)**
Potential Trainee (SCODE 4)	1	25% (49/198)	(0/0)	25% (49/198)
	2	5% (9/198)	(0/0)	5% (9/198)
	3	14% (27/198)	(0/0)	14% (27/198)
	4	5% (10/198)	(0/0)	5% (10/198)
	5	0% (0/198)	(0/0)	0% (0/198)
	Total	**48% (95/198)**	**(0/0)**	**48% (95/198)**

Table 2.6 Results of Matching to the DRF

Status	Matching Category	Number of Matches Found (% of Total Possible)		
		Successful Applications	Unsuccessful Applications	Total
Principal Investigator (SCODE 1)	1	75% (3/4)	50% (8/16)	55% (11/20)
	2	0% (0/4)	0% (0/16)	0% (0/20)
	3	0% (0/4)	0% (0/16)	0% (0/20)
	4	0% (0/4)	0% (0/16)	0% (0/20)
	5	0% (0/4)	0% (0/16)	0% (0/20)
	Total	**75% (3/4)**	**50% (8/16)**	**55% (11/20)**
Senior Investigator (SCODE 2)	1	75% (21/31)	41% (150/366)	43% (171/397)
	2	0% (0/31)	3% (11/366)	3% (11/397)
	3	4% (1/31)	9% (33/366)	9% (34/397)
	4	0% (0/31)	1% (2/366)	1% (2/397)
	5	0% (0/31)	0% (0/366)	0% (0/397)
	Total	**79% (22/31)**	**54% (196/366)**	**55% (220/397)**
Trainee (SCODE 3)	1	0% (0/32)	2% (4/220)	2% (4/252)
	2	0% (0/32)	1% (2/220)	1% (2/252)
	3	3% (1/32)	21% (47/220)	19% (48/252)
	4	31% (10/32)	7% (16/220)	10% (26/252)
	5	0% (0/32)	3% (7/220)	3% (7/252)
	Total	**34% (11/32)**	**35% (76/220)**	**34% (87/252)**
Potential Trainee (SCODE 4)	1	14% (28/198)	(0/0)	14% (28/198)
	2	3% (5/198)	(0/0)	3% (5/198)
	3	5% (10/198)	(0/0)	5% (10/198)
	4	2% (3/198)	(0/0)	2% (3/198)
	5	0% (0/198)	(0/0)	0% (0/198)
	Total	**23% (46/198)**	**(0/0)**	**23% (46/201)**

2.2.1.3 Matching to American Heart Association Grant Files

In all, matching was performed on three **AHA** grant award files. The first **AHA** file contains successful **AHA** Research Program applicants through 1990. The second and third **AHA** files contain updates for the years 1990-91 and 1992-93, respectively. For all three files, the matching methodology described in Section 2.2.1.1 was employed to identify the appropriate records.[1] The matching was performed separately for each file.

In general, a smaller percentage of individuals in the person-based data sets were matched to records in the **AHA** files than were matched to records in the **INDEX** file. For example, only 30 percent of the principal investigators matched to records in the first **AHA** file and no principal investigator matched to records in the other two **AHA** files. Table 2.7 presents the matching results for each **AHA** file by status code (SCODE) and matching category.

2.2.1.4 Matching to Medline

The methodology for matching to **Medline** differed slightly from that presented in Section 2.2.1.1. One difference stemmed from the fact that only first and last names could be used to search in **Medline**. That is, only the criteria in Matching Category Four or Matching Category Five were used (see Table 2.4). Another difference was the extremely large number of potential matches returned from the initial search, which made it impossible to hand-verify these matches.

The National Library of Medicine (NLM), using a file developed by Battelle, searched **Medline** for articles written by people in the **PMBPERS** file based upon variations of the person's first and last name. An ASCII data set containing abstracts from medical and scientific journals obtained in this search was created and given to Battelle. This information was then transformed into a SAS data set containing one record for each publication for each person. After the initial match had been performed and the data set transformed into a workable format, 107,704 records needed to be verified for accuracy. It was determined

---

[1] Date of Birth was not available in the **AHA** databases, and the matching was performed using only the criteria that did not contain Date of Birth.

Table 2.7 Matching Results for the Three AHA Files

Status	Matching Category	Number of Matches Found in the First AHA File (% of Total Possible)			Number of Matches Found In the Second AHA File (% of Total Possible)			Number of Matches Found In the Third AHA File (% of Total Possible)		
		Awarded Applications	Unsuccessful Applications	Total	Awarded Applications	Unsuccessful Applications	Total	Awarded Applications	Unsuccessful Applications	Total
Principal Investigator (SCODE 1)	2	50% (2/4)	13% (2/16)	20% (4/20)	0% (0/4)	0% (0/16)	0% (0/20)	0% (0/4)	6% (1/16)	5% (1/20)
	3	0% (0/4)	0% (0/16)	0% (0/20)	0% (0/4)	0% (0/16)	0% (0/20)	0% (0/4)	0% (0/16)	0% (0/20)
	4	0% (0/4)	13% (2/16)	10% (2/20)	0% (0/4)	0% (0/16)	0% (0/20)	0% (0/4)	0% (0/16)	0% (0/20)
	5	0% (0/4)	0% (0/16)	0% (0/20)	0% (0/4)	0% (0/16)	0% (0/20)	0% (0/4)	0% (0/16)	0% (0/20)
	Total	50% (2/4)	25% (4/16)	30% (6/20)	0% (0/4)	0% (0/16)	0% (0/20)	0% (0/4)	6% (1/16)	5% (1/20)
Senior Investigator (SCODE 2)	2	13% (4/31)	20% (74/366)	20% (78/397)	4% (1/31)	6% (23/366)	6% (24/397)	4% (1/31)	4% (15/366)	4% (16/397)
	3	0% (0/31)	1% (3/366)	1% (3/397)	0% (0/31)	0% (0/366)	0% (0/397)	0% (0/31)	1% (2/366)	1% (2/397)
	4	0% (0/31)	9% (32/366)	8% (32/397)	0% (0/31)	3% (10/366)	3% (10/397)	0% (0/31)	2% (8/366)	2% (8/397)
	5	0% (0/31)	0% (0/366)	0% (0/397)	0% (0/31)	0% (0/366)	0% (0/397)	0% (0/31)	0% (0/366)	0% (0/397)
	Total	13% (4/31)	30% (109/366)	29% (113/397)	4% (1/31)	9% (33/366)	9% (34/397)	4% (1/31)	7% (25/366)	7% (26/397)
Trainee (SCODE 3)	2	0% (0/32)	2% (5/220)	2% (5/252)	0% (0/32)	0% (1/220)	0% (1/252)	0% (0/32)	0% (1/220)	0% (1/252)
	3	0% (0/32)	0% (1/220)	0% (1/252)	0% (0/32)	0% (0/220)	0% (0/252)	0% (0/32)	0% (1/220)	0% (1/252)
	4	9% (3/32)	14% (30/220)	13% (33/252)	3% (1/32)	7% (16/220)	7% (17/252)	0% (0/32)	5% (11/220)	4% (11/252)
	5	0% (0/32)	3% (6/220)	2% (6/252)	0% (0/32)	0% (1/220)	0% (1/252)	0% (0/32)	2% (4/220)	2% (4/252)
	Total	9% (3/32)	19% (42/220)	18% (45/252)	3% (1/32)	8% (18/220)	8% (19/252)	0% (0/32)	8% (18/220)	7% (18/252)
Potential Trainee (SCODE 4)	2	2% (3/198)	0% (0/0)	2% (3/198)	2% (3/198)	0% (0/0)	2% (3/198)	2% (3/198)	0% (0/0)	2% (3/198)
	3	0% (0/198)	0% (0/0)	0% (0/198)	0% (0/198)	0% (0/0)	0% (0/198)	1% (1/198)	0% (0/0)	1% (1/198)
	4	3% (5/198)	0% (0/0)	3% (5/198)	0% (0/198)	0% (0/0)	0% (0/198)	1% (1/198)	0% (0/0)	1% (1/198)
	5	0% (0/198)	0% (0/0)	0% (0/198)	0% (0/198)	0% (0/0)	0% (0/198)	1% (2/198)	0% (0/0)	1% (2/198)
	Total	4% (8/198)	0% (0/0)	4% (8/198)	2% (3/198)	0% (0/0)	2% (3/198)	4% (7/198)	0% (0/0)	4% (7/198)

that verifying these matches by hand would not be cost-effective. Thus, an additional computer algorithm was developed to eliminate erroneous matches. This process is described below.

The process began by extracting those records that were the only matches for a specified author from the POEMB person file. Of the 107,704 records, there were 23 records that were the only match for the given author, leaving 107,681 records with multiple matches per author.

Next, the multiple matches that did not cite NIH support were removed (NIH support was determined by scanning the **Medline** ID field for two-character grant abbreviations that signified that the institution belonged to the NIH). Of the 107,681 records, NIH support was cited in only 35,859 records. Finally, the 35,859 records that cited NIH support were checked to determine whether they contained one of the grant numbers in the final **CGAF** file (see Section 2.2.1.2). Overall, there were 9,001 records in the **Medline** file that had a grant number that also appeared in the **CGAF** file. These 9,001 records were then combined with the 23 single match records for a total of 9,024 records in the final **Medline** analysis file. Table 2.8 presents the results of the matching for each status code (SCODE).

Table 2.8 Matching Results for Medline

Status (SCODE)	Number of People With Articles	Number of Articles
Principal Investigator (SCODE 1)	18	489
Senior Investigator (SCODE 2)	350	7,220
Potential Trainee (SCODE 3)	158	743
Trainee Cited in Progress Report (SCODE 4)	98	572
Total	624	9,024

2.2.2 Creation of the Analysis Data Sets

Section 2.2.1 above describes the matching performed on six databases. Each matching activity resulted in a data set. For the **TFF** and **DRF**, no further manipulations were performed, and these files were used as is in the analysis. However, the resulting databases for **CGAF**, **CRISP**, **AHA**, and **Medline** were further modified to enhance the statistical analysis. The following sections summarize each data set and the modifications, if any, that were made.

2.2.2.1 Trainee and Fellow File

The **Trainee and Fellow File (TFF)** contains records of fellowships and traineeships awarded to researchers named in POEMB applications or progress reports. The **TFF** analysis file was matched to selected information from the **PMBPERS** data set to produce a data set that represented all fellowships and traineeships awarded to POEMB trainees as of 1994 (some individuals were awarded more than one traineeship or fellowship). This data set was used for the analyses described in Chapter 3.

2.2.2.2 Doctorate Records File

The **Doctorate Records File (DRF)** contains records of all doctorates earned by researchers named in POEMB applications or progress reports granted by regionally accredited United States universities, in all fields, from 1920 to 1993. The **DRF** analysis file was matched to selected information from the **PMBPERS** data set to produce a data set that represented all doctorates earned by POEMB trainees as of 1993 (some individuals earned more than one doctorate). This data set was used for the analyses described in Chapter 3.

2.2.2.3 Consolidated Grant Applicant File and Computer Retrieval of Information on Scientific Projects

The **CGAF/CRISP** analysis file combines records of grant applications (**CGAF**) and grant awards (**CRISP**) made *to* (applications) or *by* (awards) NIH and other Public Health Service organizations for the years 1938 to 1993. Only the names of the primary investigators (PIs) associated with each grant application (and not support staff) were available.

The **CGAF/CRISP** analysis file was based upon the files resulting from the matching discussed in Section 2.2.1. However, additional modifications were made to the file to facilitate the statistical analysis.

First, all records pertaining to contract activity were removed. Next, multiple entries generated by amendments to the original grant application and competing continuations were removed from the file so that grant activity would be represented only by new applications. Finally, records corresponding to an

investigator's first POEMB application were also removed from the **CGAF/CRISP** analysis file, as were grants awarded more than four years prior to an individual's first POEMB application.

Additional information was derived from the **PMBAPP** file and included in the **CGAF/CRISP** analysis file. This information included the date of an individual's first POEMB application (ZDATE) and a variable to indicate whether that application was awarded (PAWARD). The second POEMB application for those individuals that were on more than one application was not deleted, and it is indicated in the variable POEMB.

We then created a variable named QUARTER to indicate the quarter in which the grant application was submitted relative to the first POEMB application. **CGAF** grant applications were assigned to quarters relative to each person's first POEMB application. Quarter "0" was defined so that a person's first POEMB application occurs at the midpoint of the quarter.

Records with missing grant information were added for person and quarter combinations where no grant information existed (i.e., the individual was not a principal investigator on a NIH/PHS grant during that quarter). An indicator of these added records is included in the baseline file (INCGAF). This gives a data set with at least one record for each person by quarter combination. (See Table 2A.2 of Appendix 2A for a list and definitions of the variables contained in the **CGAF/CRISP** analysis file).

The Institute/Center/Division (ICD) code in **CGAF** was used to classify grants by institution. Grants were classified as heart/lung (HL)-related if the ICD code indicated that the awarding institution was the National Heart, Lung, and Blood Institute (NHLBI) or one of its divisions. For the purposes of this analysis, all other grant applications were considered non-heart/lung-related (non-HL). Table 2A.5 in Appendix 2A lists the seven ICD codes that were classified as HL.

CRISP descriptors (or subject terms) were used to classify the nature of the NIH/PHS awards received by successful and unsuccessful POEMB applicants during the period 17 quarters before and 16 quarters after POEMB application. First, a list was compiled of all the **CRISP** subject terms associated with the awards received by individuals in both POEMB applicant groups. This list was alphabetized and manually checked using the ***CRISP Intramural Research Index*** to verify the classification of terms. This index contains a complete listing of **CRISP** subject headings used to describe research projects funded by NIH. Awards with **CRISP** descriptors listed under the major headings "cardiac," "cardiovascular," "heart," and "respiratory" were classified as heart/lung (HL). Awards with descriptors listed under the major heading "molecular biology" were classified as molecular biology (MB). Awards with all other descriptors were classified as "other" (CC). A complete list of **CRISP** descriptors that fall within the

categories of interest are presented in Appendix 2A. Because each award is associated with one or more descriptor, there are more than three possible **CRISP** classification categories. These categories include: HL only, MB only, HL/MB (mixed), "other" only, HL/ "other," MB/ "other," and HL/MB/ "other." The **CGAF/CRISP** file was used for the analyses described in Chapter 4.

2.2.2.4 American Heart Association Grant Files

The **AHA** analysis file contains information from the three American Heart Association (AHA) files (**AHA1**, **AHA2**, **AHA3**), the **PMBAPP** file, and the **PMBPERS** file. These files are subsets of larger **AHA** data sets and were obtained by matching the **AHA** files to the **PMBPERS** file. Only those records associated with a POEMB applicant were included in these subsets. The **AHA1** file contains successful AHA Research Program Applicants through 1990. The **AHA2** and **AHA3** files contain updates to **AHA1** for the years 1990-91 and 1992-93, respectively. AHA applications were designated as occurring before (-1), during (0), or after (1) an individual's first POEMB application (B4AFTR). The value for this variable was determined by comparing the year of the AHA application to the year of POEMB application. Records with missing AHA information were added for person and B4AFTR (-1 to 1) combinations where no AHA information existed (i.e., the individual did not receive an AHA award). An indicator variable was included (AAWARD) to distinguish records that were awards (1) from those that were not (0), thus generating a data set with at least one record for each person by B4AFTR combination. Table 2A.3 in Appendix 2A contains a complete listing of all variables included in the **AHA** baseline file.

The **AHA** file was used for the analyses described in Chapter 4.

2.2.2.5 Medline

The **Medline** analysis file contains information from the **Medline** file, the **PMBAPP** file, and the **PMBPERS** file. New variables were derived from the **PMBAPP** file. The date of a person's first POEMB application (ZDATE) was derived and included. A variable was added to indicate whether a person's first POEMB application was awarded (PAWARD). The year and month of publication (YMPUB) were derived from the variable DATEPUB, which contained date of publication information in an inconsistent format. Articles were assigned to quarters (QUARTER) relative to the person's first POEMB application. Quarter 0 was defined such that a person's first POEMB application occurs at the

midpoint of the quarter. Quarter -1 would be the quarter immediately preceding this quarter and quarter 1 would be the following quarter, etc.

Using the journal title (TA), journals were classified in variable JCAT as Blood (B), Heart (H), Heart/Blood (H/B), Heart/Lung (H/L), Heart/Molecular Biology (H/M), Lung (L), Lung/Molecular Biology (L/M), Molecular Biology (M), Other (O), or Cross-Cutting (X). MeSH terms were used to create indicators to show whether or not an article discussed issues related to Heart (HI), Heart/Lung (HLI), Heart/Molecular Biology (HMI), Lung (LI), Molecular Biology (MI), and Other (OI).

Records with missing **Medline** information were added for person and QUARTER (-16 to 16) combinations where no **Medline** information existed (i.e., the person did not publish an article). An indicator is included (INMED) to distinguish records that were publications (1) and those that were not (0). This gives a data set with at least one record for each person by QUARTER combination.

A complete listing of all variables in the **Medline** analysis file is in Table 2A.4 in Appendix 2A. The variable, its origin, and a brief description are included. More detailed descriptions of individual variables can be found in the documentation of the file of origin.

The **Medline** file was used for the analyses described in Chapter 5.

3.0 Training of Investigators and Trainees

3.0 Training of Investigators and Trainees

The analysis in this chapter is intended to characterize the Programs of Excellence in Molecular Biology (POEMB) grant program of the National Heart, Lung, and Blood Institute (NHLBI) in terms of its senior investigators and trainees. To accomplish this Battelle has:

- Identified the area of doctoral-level training for investigators cited on POEMB applications and trainees cited in POEMB progress reports.
- Determined whether POEMB training is a substitute for or in addition to other NIH awards.

In this chapter we describe the data and methodology used to complete the above tasks and then present the major findings. Section 3.1 contains a discussion of areas of doctoral level training. Section 3.2 contains a discussion of other awards received by POEMB trainees.

3.1 Areas of Doctoral-level Training

This section begins by describing the data used for analysis (Section 3.1.1) and concludes with a discussion of methodology and findings.

3.1.1 Data

To determine the areas of doctoral-level training for researchers cited in POEMB applications and progress reports, we merged the Battelle-developed **PMBPERS** file with a subset from the Doctorate Records File (**DRF**). (See Appendix 2A for a summary of fields in **PMBPERS** and Appendix 3A for a summary of fields used from **DRF**.) The **PMBPERS** file consisted of 871 records, one for each researcher appearing on a POEMB application or progress report. The **DRF** subset was made up of 382 records belonging to those researchers who had earned a doctorate degree and were cited in POEMB applications and progress reports. The merge resulted in a data set containing 877 records (certain individuals had more than one earned doctorate recorded in the **DRF**). For details on matching procedures see Section 2.2 of this report.

Table 3.1 at the end of this chapter lists the number of researchers that appear on POEMB applications or progress reports by their insititution and level of responsibility. Trainees (SCODE=4) will only appear for the three successful applicants, since only winning institutions could request support for trainees. Frequency tables were constructed using ICODE, SCODE, DEG1, and DEG2 from **PMBPERS**, and PHDFIELD from **DRF**. A more detailed description of the methodology and a preliminary report of the findings can be found below.

3.1.2 Methodology and Findings

Doctorate Degrees Earned. Since the PHDFIELD variable from the **DRF** file does not contain information about professional doctorates, we used DEG1 and DEG2 from the Battelle **PMBPERS** file to determine what types of doctorate degrees were earned by researchers.[1] Table 3.2 at the end of this chapter presents the types of doctorate degrees earned by senior investigators only (SCODE=1 or SCODE=2). Doctorate degrees earned by POEMB trainees are shown in Figure 3.1.

The results in Table 3.2, at the end of this chapter, were stratified to compare successful and unsuccessful applicant institutions in terms of their mix of professional and academic doctorates. Figure 3.2 shows the results of this comparison and suggests there was a smaller percentage of MDs among the winning applicants than among the losing applicants.

[1] Data in the **PMBPERS** file were extracted from POEMB applications and progress reports and reflect the information as provided by each applicant institution.

Figr 3.1

Successful vs Unsuccessful POEMB Applicants
Doctorate Degrees of Senior Investigators

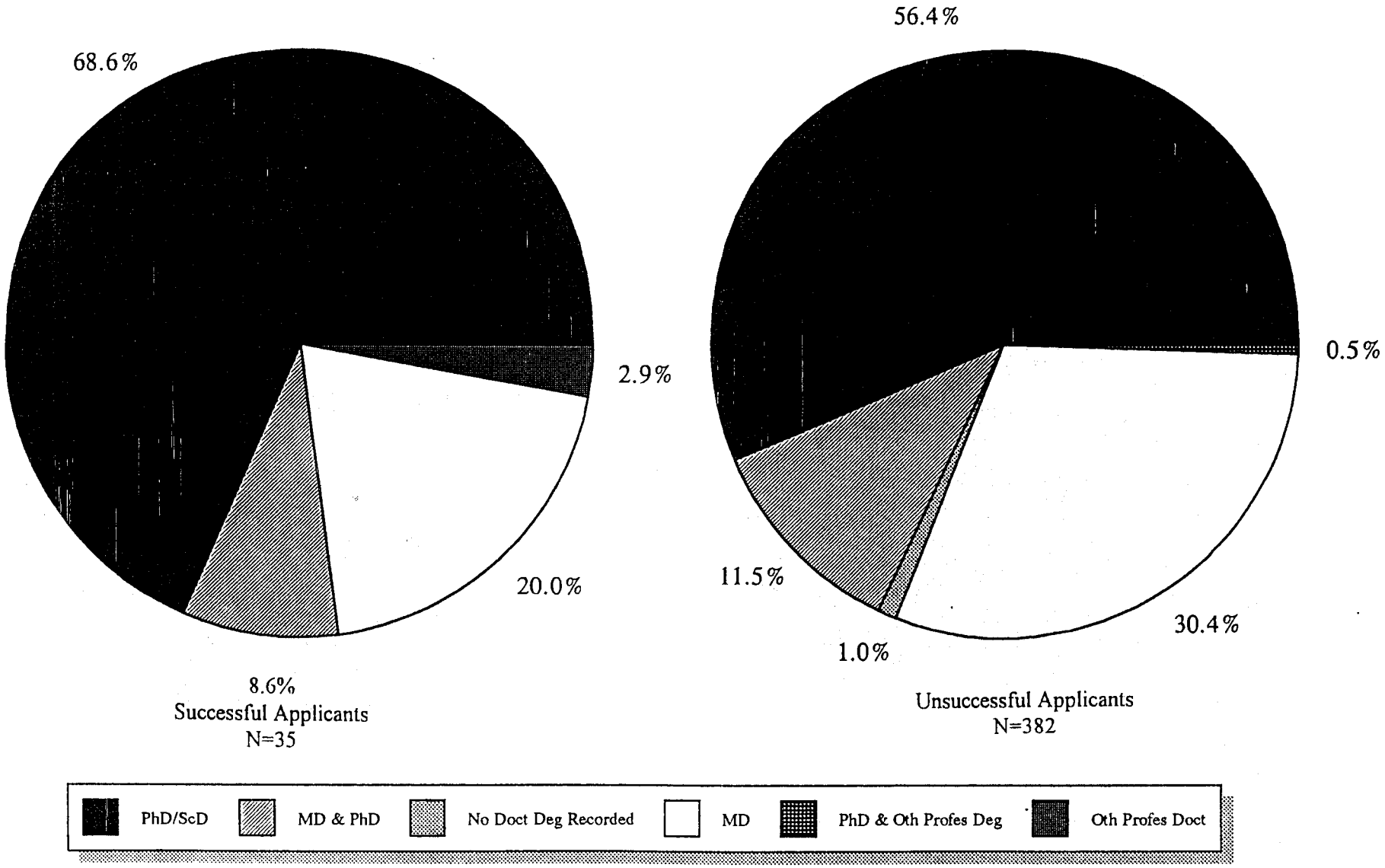
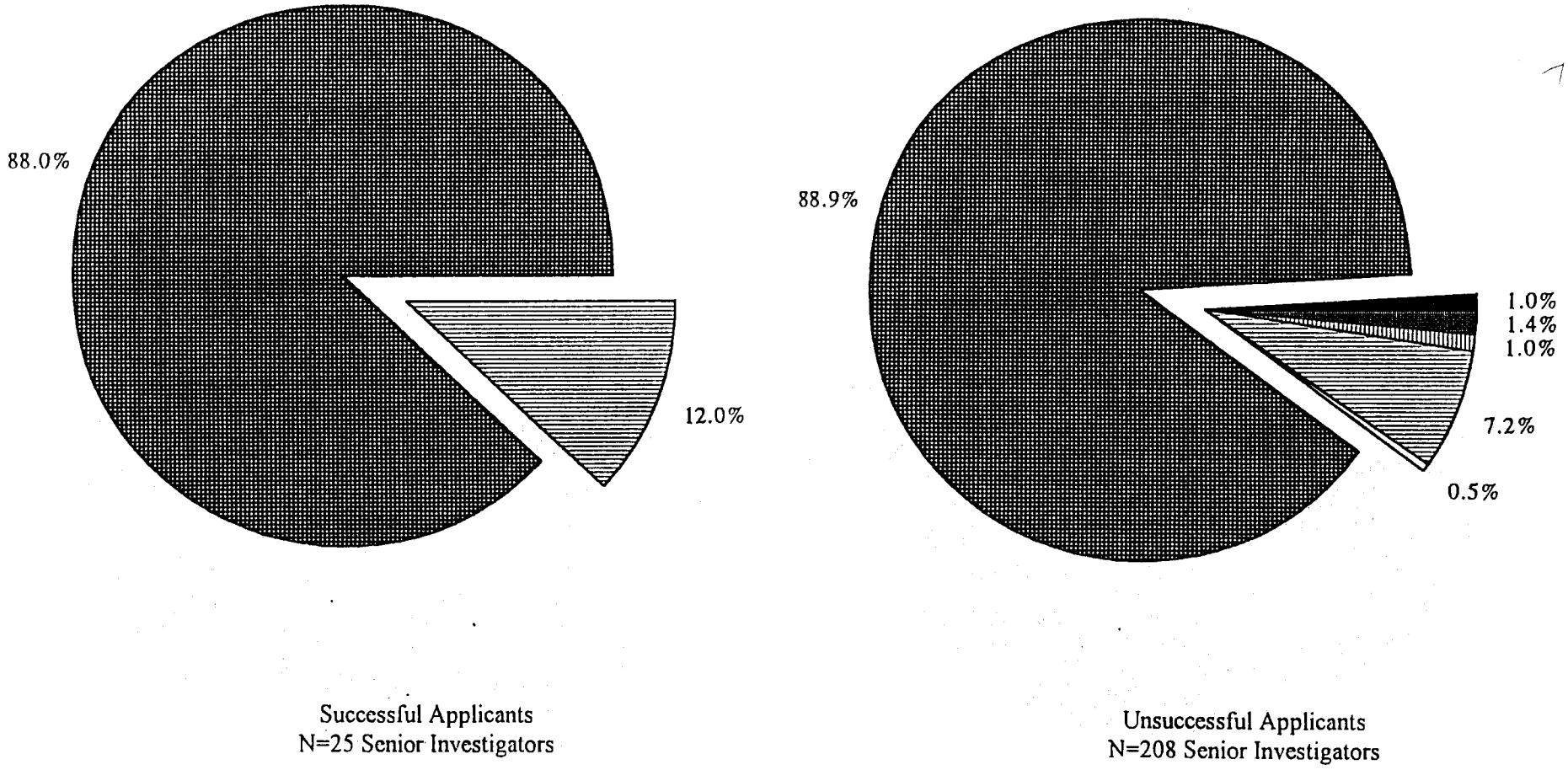


Figure 3.2

Successful vs Unsuccessful POEMB Applicants
Senior Investigators' Major Fields of PhDs or ScDs



Successful Applicants
N=25 Senior Investigators

Unsuccessful Applicants
N=208 Senior Investigators

Disciplinary Fields of PhDs or ScDs. After evaluating the types of doctorates earned by the researchers, we then focused on the individuals who had earned an academic (as opposed to a professional) doctorate. We used the variable PHDFIELD from the **DRF** to describe the disciplinary fields of study. As seen in Table 3.3 at the end of this chapter and in Figure 3.3, senior investigators and trainees were overwhelmingly trained in the Biological Sciences (PHDFIELD codes 100-199). A breakdown of the various biological sciences is provided in Tables 3.4 and 3.5 at the end of this chapter for the senior investigators and the trainees, respectively. We compare successful and unsuccessful applicant institutions in terms of the mix of disciplinary fields. Figures 3.4 through 3.6 show that the successful and unsuccessful applicant instructions had a very similar profile in this respect.

3.2 Other Awards Received by POEMB Trainees

This section begins by describing the data used for analysis (Section 3.2.1) and concludes with a discussion of methodology and findings (Section 3.2.2).

3.2.1 Data

Below we discuss data used to identify (1) other training grants awarded to POEMB trainees and (2) research grants and contracts awarded to POEMB trainees.[2]

3.2.1.1 Data to Identify Other Training Grants Awarded to POEMB Trainees

To determine whether POEMB training was instead of or in addition to other training grants, we merged a subset from the **PMBPERS** file with a subset of the **TFF**. (See Appendix 3A for a summary of fields used from the **TFF**.) The **PMBPERS** file consisted of 266 records, one for each trainee (status 4) that appeared on a POEMB progress report. The **TFF** subset (N = 1929) contained records of fellowships and traineeships awarded (AWARD INDICATOR = 1) to researchers named in POEMB applications or

[2] Further detail on preparation of the analysis data sets is provided in Section 2.2 of this report.

Figure 3.3

Unsuccessful POEMB Applicants: Areas of Concentration for Senior Investigators with PhDs or ScDs in the Biological Sciences

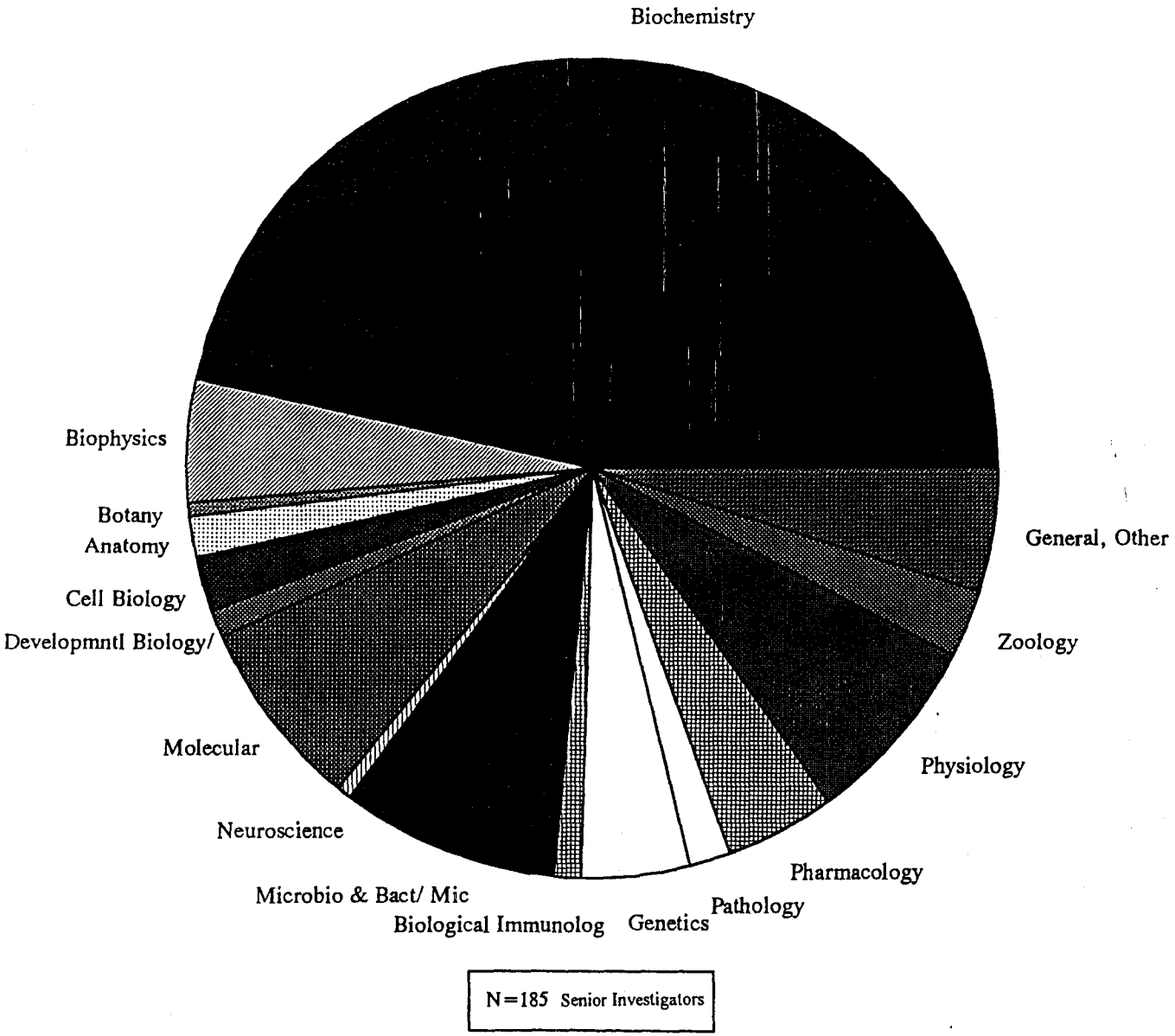


Fig 3.4

Successful POEMB Applicants: Areas of Concentration for Senior Investigators with PhDs or ScDs in the Biological Sciences

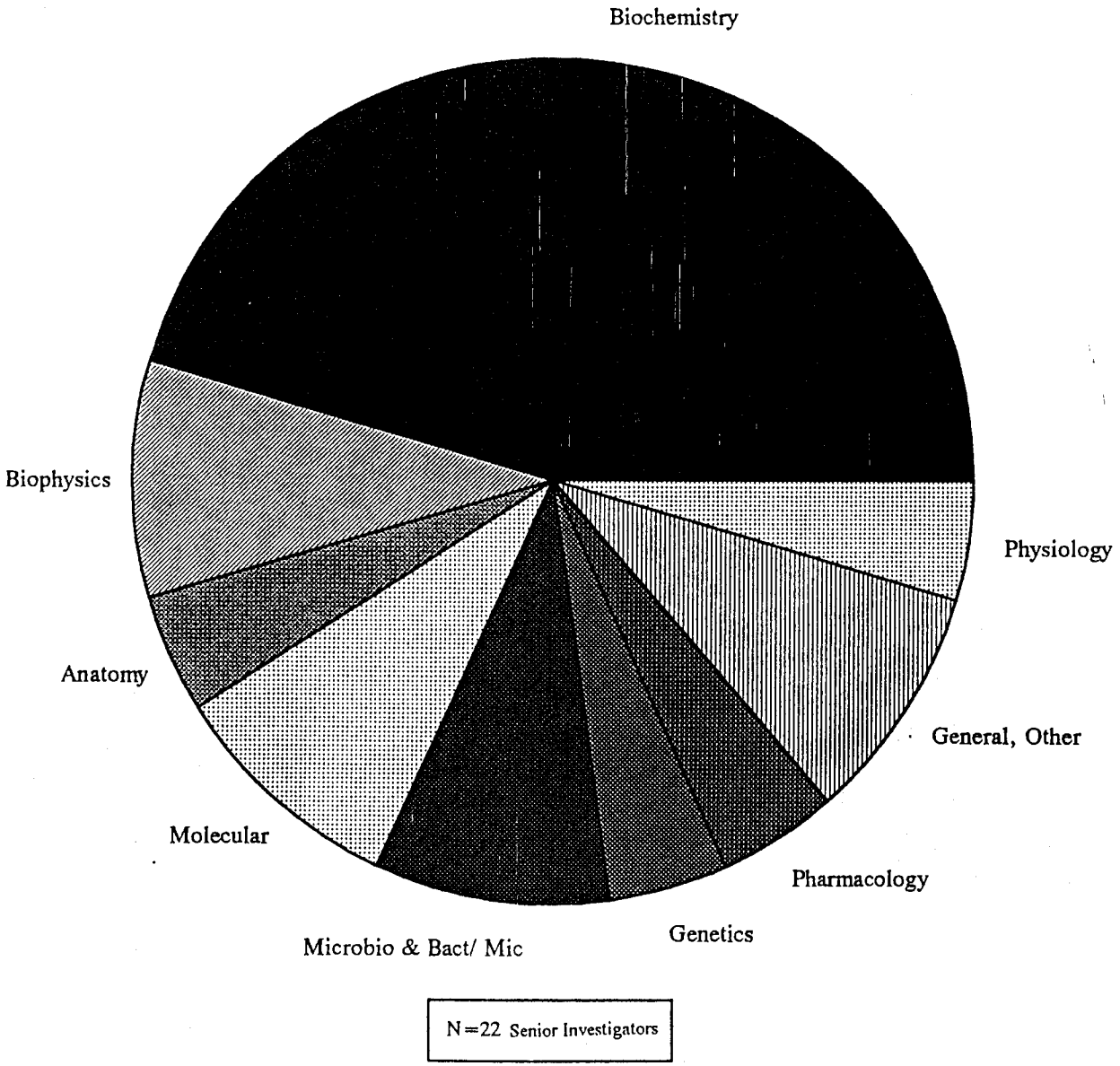


Figure 3.5

Successful POEMB Applicants
Doctorate Degrees of Trainees

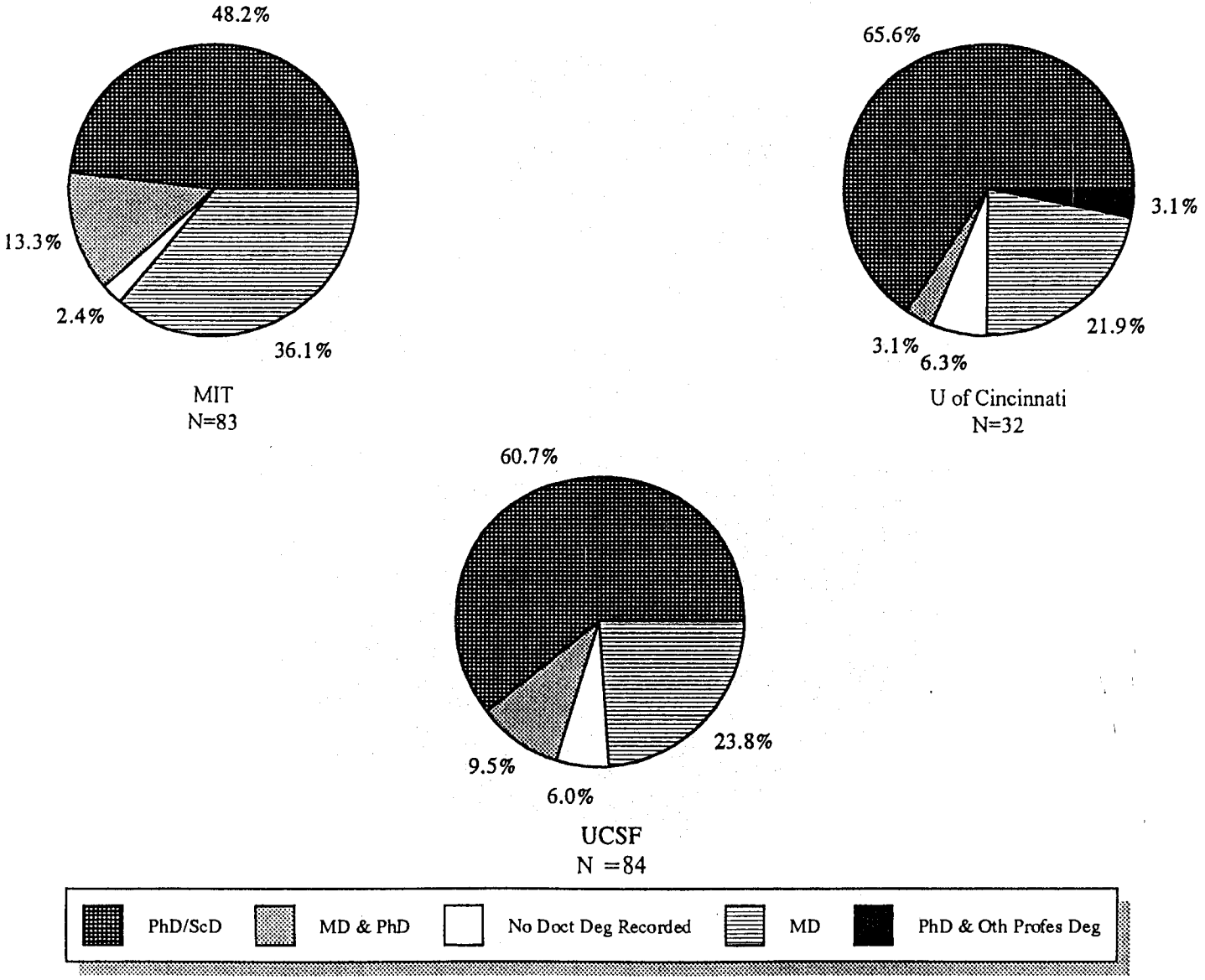
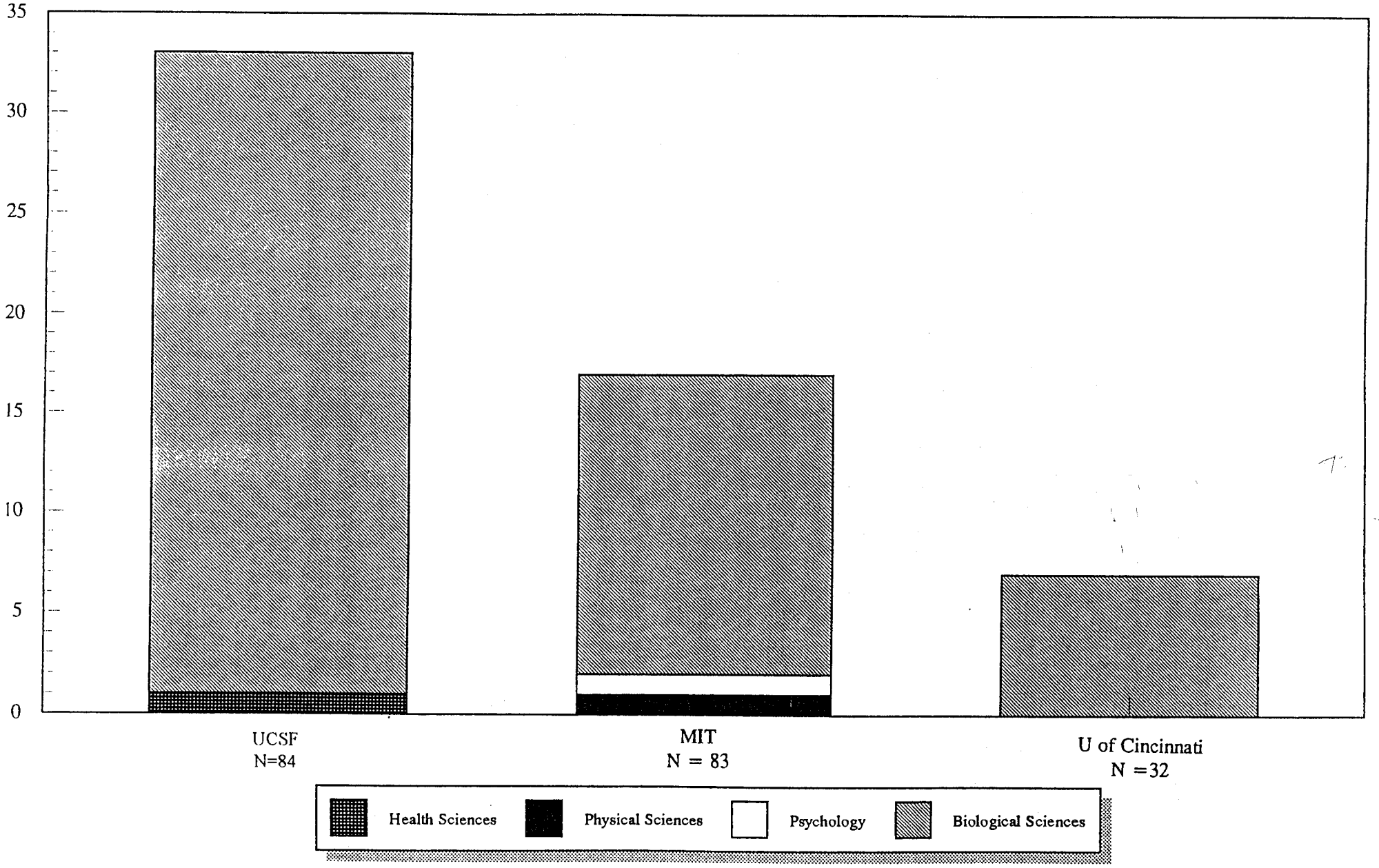


Figure 3.6

Successful . EMB Applicants

Trainees' Major Fields of PhDs or ScDs



progress reports. Merging the two subsets produced a data set with 275 records, representing all fellowships and traineeships awarded to POEMB trainees as of 1994 (some individuals were awarded more than one traineeship or fellowship). The variables ICODE, PCODE, and the progress report status variables (_1 , _2, _3, _4, _5, _6, and _7) from Battelle's **PMBPERS** file, as well as the variables ACTIVITY/PROGRAM CODE, PHS ID NUMBER, TENURE-ENTRY DATE and TENURE-EXIT DATE from the **TFF** were used for this component of the analysis. A more descriptive account of the methodology and summary tables and figures are provided in Section 3.2.2.

3.2.1.2 Data to Identify Research Grants and Contracts Awarded to POEMB Trainees

To identify other research grants and contracts awarded to POEMB trainees, we merged the subset of the **PMBPERS** file with a subset of the **CGAF**. Again, the **PMBPERS** file consisted of 266 records, one for each trainee that appeared on a POEMB progress report. The **CGAF** subset (N=10,976) contained records grants and contracts awarded (Award Indicator=1) to researchers named in POEMB applications or progress reports. Merging the two subsets produced a data set with 303 records, representing all grants and contracts awarded to POEMB trainees as of 1994 (again some trainees were awarded more than one grant or contract). The variables ICODE, PCODE, _1, _2, _3, _4, _5, _6, _7 from **PMBPERS**, and the variables ACTIVITY PROGRAM CODE, INSTITUTE (NIH) IDENTIFICATION NUMBER, AWARD INDICATOR, START DATE 1, END DATE1 from **CGAF** were used for this component of the analysis. (See Appendix 3A for a summary of fields used from the **CGAF**.) Section 3.2.2 provides a more detailed discussion of the methodology and presents tables and figures that summarize the findings.

3.2.2 Methodology and Findings

Below we present methodology and findings relating to (1) other training grants awarded to POEMB trainees and (2) other research grants and contracts awarded to POEMB trainees.

3.2.2.1 Other Training Grants Awarded to POEMB Trainees

Figure 3.7 shows the number of POEMB trainees receiving other types of traineeships and fellowships. We used the variables _1, _2, _3, _4, _5, _6, and _7 from Battelle's **PMBPERS** file, as well as the variables ACTIVITY, TENURE-ENTRY DATE, and TENURE-TERMINATION DATE from the **TFF** to identify the time sequence in which awards were received. To determine the first year of POEMB support, we applied the following algorithm:

For trainees from Massachusetts Institute of Technology and the University of Cincinnati:

- If _1 was checked with an 'X', then the first year of POEMB support began on October 1988 which was the date of the first progress report.
- If _1 was not checked with an 'X' but _2 had a value of 'X', then the first year of POEMB support began on October 1989, which was the date of the second progress report.
- If _1 and _2 were blank and _3 was checked with an 'X', then October 1990 was the first year of POEMB support.
- If _1, _2, and _3 were blank and _4 was checked with an 'X', then the first year of POEMB support was October 1991, the date of the fourth progress report.
- If _1, _2, _3, and _4 were blank while _5 was checked with an 'X', then the first year of POEMB support was October 1992, the date of the fifth progress report.
- If _1, _2, _3, _4, and _5 were left blank while _6 was assigned a value of 'X', then the first year of POEMB support was October 1993, the date of the sixth progress report.
- If _7 was the only _N variable checked with an 'X', then the first year of POEMB support was October 1994, the date of the seventh and final progress report.

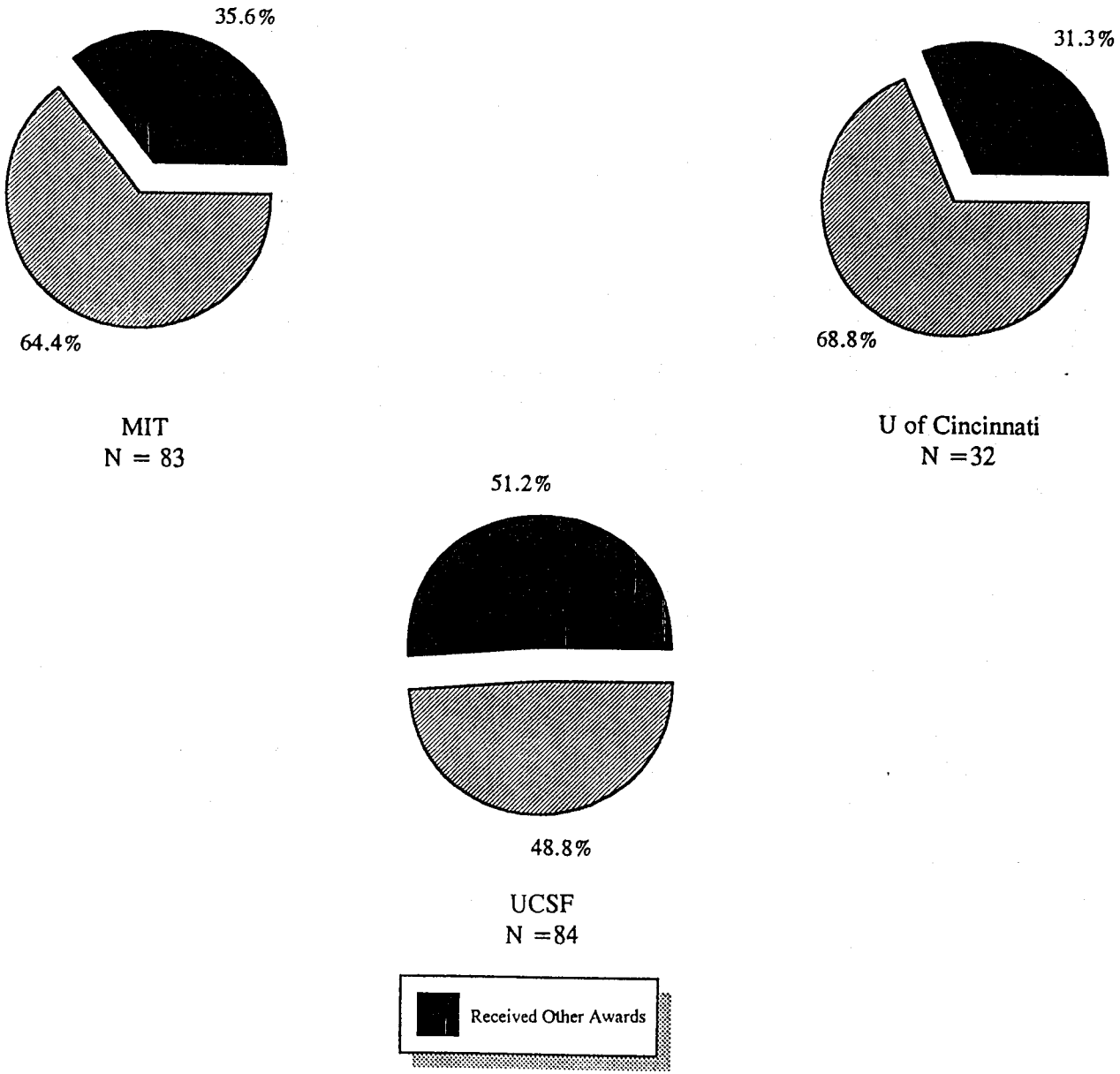
For trainees from the University of California at San Francisco:

- If _1 was checked with an 'X', then the first year of POEMB support began on October 1989, the date of the first progress report.[3]

[3] Note: the University of California/San Francisco commenced its POEMB grant one year later than did the other two grantees, 1989 as opposed to 1988.

Figure 3.7

Percentage of POEMB Trainees with Other Types of Training Awards



- If _1 was not checked with an 'X' but _2 had a value of 'X', then the first year of POEMB support began on October 1990, the date of the second progress report.
- If _1 and _2 were blank and _3 was checked with an 'X', then October 1991 was the first year of POEMB support.
- If _1, _2, and _3 were blank and _4 was checked with an 'X', then the first year of POEMB support was October 1992, the date of the fourth progress report.
- If _1, _2, _3, and _4 were blank while _5 was checked with an 'X', then the first year of POEMB support was October 1993, the date of the fifth progress report.
- If _1, _2, _3, _4, and _5 were left blank while _6 was assigned a value of 'X', then the first year of POEMB support was October 1994, the date of the sixth and latest progress report.

Using the **TFF** variable ACTIVITY, we then describe which types of traineeships and fellowships were awarded to POEMB trainees. Tables 3.6, 3.7, and 3.8 at the end of this chapter list the types of awards that trainees received before, during, and after POEMB support, respectively. Because some trainees received more than one award, the total number of traineeships and fellowships is greater than the total number of trainees.

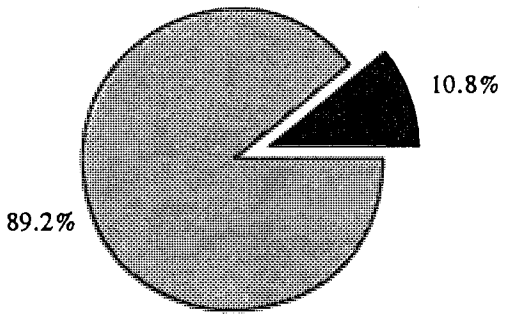
The tables show that POEMB trainees were recipients of other traineeships and fellowships prior to receiving POEMB support. During their POEMB traineeships, some trainees received additional support, namely through F05 International Fellowships, F31 Postdoctoral Fellowships, F32 Postdoctoral Individual NRSAs, and the T32 Institutional National Research Service Award. Some trainees ended their affiliation with POEMB and went on to receive F32 Postdoctoral Individual NRSA awards or T32 Institutional National Research Service Awards.

3.2.2.2 Other Research Grants and Contracts Awarded to POEMB Trainees

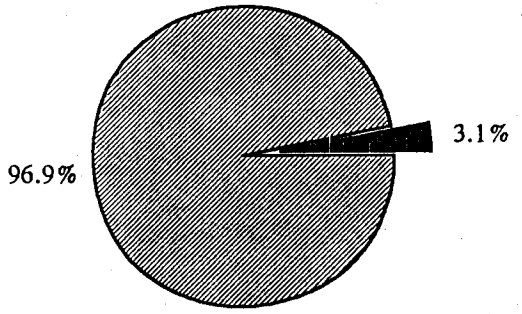
Figure 3.8 presents the number of POEMB trainees receiving other types of grants and contracts. We used the variables _1, _2, _3, _4, _5, _6, and _7 from the **PMBPERS** file, and the variables START DATE1, and END DATE1 from **CGAF** to identify the time sequence in which awards were received. To determine the first year of POEMB support, we applied the same algorithm described in the preceding section.

Figure 3.8

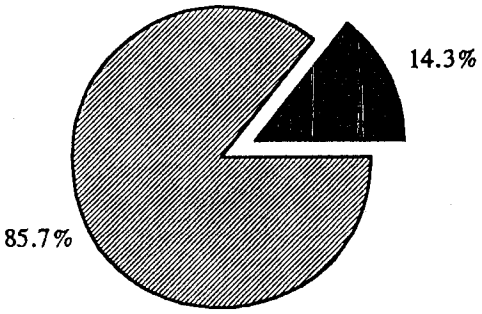
Percentage of POEMB Trainees with Other Research Grants or Contracts



MIT
N = 83



U of Cincinnati
N =32



UCSF
N= 84

Received Other Awards

Using the variable ACTIVITY from **CGAF**, we then describe which types of grants and contracts were awarded to POEMB trainees. Tables 3.9, 3.10, and 3.11 at the end of this chapter list the types of awards that trainees received before, during, and after POEMB support, respectively. Again, because some trainees received more than one award, the total number of grants and contracts is greater than the total number of trainees.

The tables suggest that there were some trainees that received research grants or contracts from NIH and other Public Health Service organizations prior to participation in POEMB. During the period of POEMB support, some trainees received other research grants and contracts, most commonly through the K11 Physician Scientist Award, the R01 Research Projects Grants (Traditional), and the T32 Institutional National Research Service Award. Certain individuals have ended their stint with POEMB and have since received grants or contracts through the R01 Research Projects Grant (Traditional), the R03 Small Research Grant, or the R29 First Independent Research Support and Transition (FIRST) Award.

Table 3.1 All POEMB Applicants: Research Teams by Levels of Responsibility

APPLICANTS		RESEARCH TEAMS			
Institution	Department	Senior Investigators	Other Staff	Trainees	Total
Mass Inst of Tech[a]	Bio	9 (9.8%)	0 (0.0%)	83 (90.2%)	92 (100%)
U of Cincinn[a]	Microbio	11 (14.9%)	31 (41.9%)	32 (43.2%)	74 (100%)
U of Calif San Fran[a]	Cardiovas Resear Inst	12 (12.0%)	1 (1.0%)	87 (87.0%)	100 (100%)
Wash U	Ped	18 (47.4%)	20 (52.6%)	0 (0.0%)	38 (100%)
U of Penn	Cardiovas Pulmon Div	31 (100.0%)	0 (0.0%)	0 (0.0%)	31 (100%)
U of Calif San Diego	Medicine	14 (93.3%)	1 (6.7%)	0 (0.0%)	15 (100%)
U of Chicago	Medicine	12 (42.9%)	16 (57.1%)	0 (0.0%)	28 (100%)
U of Illinois	Internal Med	15 (75.0%)	5 25.0%)	0 (0.0%)	20 (100%)
U of Ala Birmingham	Medicine	16 (31.4%)	35 (68.6%)	0 (0.0%)	51 (100%)
Wash U	Ped	14 (87.5%)	2 (12.5%)	0 (0.0%)	16 (100%)
Baylor	Cell Bio	24 (100.0%)	0 (0.0%)	0 (0.0%)	24 (100%)
Childrn's Hosp Boston	Cardiol	19 (95.0%)	1 (5.0%)	0 (0.0%)	20 (100%)
Baylor	Col of Med	24 (52.2%)	22 (47.8%)	0 (0.0%)	46 (100%)
U of Wash	Pathology	33 (100.0%)	0 (0.0%)	0 (0.0%)	33 (100%)
Johns Hopkins	Biochem	11 (50.0%)	11 (50.0%)	0 (0.0%)	22 (100%)
W. Alton Jones Cell Science Ctr	Cell Science	7 (53.8%)	6 (46.2%)	0 (0.0%)	13 (100%)
U of Illinois	Food Chem	1 (16.7%)	5 (83.3%)	0 (0.0%)	6 (100%)
U of Calif San Diego	Pharmacol	22 (78.6%)	6 (21.4%)	0 (0.0%)	28 (100%)
Duke	Med Ctr	11 (45.8%)	13 (54.2%)	0 (0.0%)	24 (100%)
U of Iowa	Cardiovas Ctr	38 (60.3%)	25 (39.7%)	0 (0.0%)	63 (100%)
U of Calif San Fran	Metabolic Resear Unit	8 (29.6%)	19 (70.4%)	0 (0.0%)	27 (100%)
Childrn's Hosp Boston	Cardiology	5 (20.8%)	19 (79.2%)	0 (0.0%)	24 (100%)
StateU of New York Syracuse	Anatomy & Cell Bio	29 (82.9%)	6 (17.1%)	0 (0.0%)	35 (100%)
Boston U	Sch of Med	30 (75.0%)	10 (25.0%)	0 (0.0%)	40 (100%)

[a] Applicants were POEMB recipients.

Table 3.2 All POEMB Applicants: Doctorate Degrees of Senior Investigators

APPLICANTS		N	DOCTORATE DEGREES [b]					
Institution	Department	Senior Invstgtrs	PhD or ScD	MD	Prof Deg other than MD	PhD plus MD	PhD plus Prof Deg other than MD	No Doct Deg Noted
Mass Inst of Tech[a]	Bio	9 (100%)	6 (67%)	1 (11%)	1 (11%)	1 (11%)	0 (0%)	0 (0%)
U of Cincinn[a]	Microbio	11 (100%)	10 (91%)	1 (9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
U of Calif San Fran[a]	Cardiovas Resear Inst	12 (100%)	8 (67%)	2 (17%)	0 (0%)	2 (17%)	0 (0%)	0 (0%)
Wash U	Ped	18 (100%)	8 (44%)	7 (39%)	0 (0%)	2 (11%)	0 (0%)	1 (6%)
U of Penn	Cardiovas Pulmon Div	31 (100%)	15 (48%)	9 (29%)	0 (0%)	5 (16%)	1 (3%)	1 (3%)
U of Calif San Diego	Medicine	14 (100%)	7 (50%)	3 (21%)	0 (0%)	4 (29%)	0 (0%)	0 (0%)
U of Chicago	Medicine	12 (100%)	7 (58%)	3 (25%)	0 (0%)	2 (17%)	0 (0%)	0 (0%)
U of Illinois	Internal Med	15 (100%)	14 (93%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	0 (0%)
U of Ala Birmingham	Medicine	16 (100%)	13 (81%)	1 (6%)	0 (0%)	2 (13%)	0 (0%)	0 (0%)
Wash U	Ped	14 (100%)	2 (14%)	8 (57%)	0 (0%)	3 (21%)	0 (0%)	1 (7%)
Baylor	Cell Bio	24 (100%)	16 (67%)	5 (21%)	0 (0%)	3 (13%)	0 (0%)	0 (0%)
Childrn's Hosp Boston	Cardiol	19 (100%)	8 (42%)	9 (47%)	0 (0%)	2 (11%)	0 (0%)	0 (0%)
Baylor	Col of Med	24 (100%)	13 (54%)	9 (38%)	0 (0%)	2 (8%)	0 (0%)	0 (0%)
U of Wash	Pathology	33 (100%)	20 (61%)	7 (21%)	0 (0%)	5 (15%)	1 (3%)	0 (0%)
Johns Hopkins	Biochem	11 (100%)	5 (45%)	2 (18%)	0 (0%)	4 (36%)	0 (0%)	0 (0%)
W. Alton Jones Cell Science Ctr	Cell Science	7 (100%)	6 86%)	0 (0%)	0 (0%)	1 (14%)	0 (0%)	0 (0%)
U of Illinois	Food Chem	1 (100%)	1 (100%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
U of Calif San Diego	Pharmacol	22 (100%)	14 (64%)	8 (36%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Duke	Med Ctr	11 (100%)	4 36%)	7 (64%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
U of Iowa	Cardiovas Ctr	38 (100%)	17 (45%)	18 (47%)	0 (0%)	3 (8%)	0 (0%)	0 (0%)
U of Calif San Fran	Metabolic Resear Unit	8 (100%)	4 50%)	2 (25%)	0 (0%)	2 (25%)	0 (0%)	0 (0%)
Childrn's Hosp Boston	Cardiology	5 (100%)	1 (20%)	2 (40%)	0 (0%)	1 (20%)	0 (0%)	1 (20%)
SUNY/ Syracuse	Anatomy & Cell Bio	29 (100%)	25 (86%)	2 (7%)	0 (0%)	2 (7%)	0 (0%)	0 (0%)
Boston U	Sch of Med	30 (100%)	16 (53%)	14 (47%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

a Applicants were POEMB recipients

b Rounding may lead to totals greater or less than 100%

Table 3.3 All POEMB Applicants: Senior Investigators' Major Fields of First PhDs or ScDs

APPLICANTS		N	MAJOR FIELDS OF FIRST PhD OR ScD		
Institution	Department	Senior investigators w/PhDs or ScDs	Biological Sciences	Physical Sciences	Other Fields[b]
Mass Inst of Tech[a]	Bio	7	7	0	0
U of Cincinn[a]	Microbio	9	8	1	0
U of Calif San Fran[a]	Cardiovas Resear Inst	9	7	2	0
Wash U	Ped	10	8	1	1
U of Penn	Cardiovas Pulmon Div	18	17	0	1
U of Calif San Diego	Medicine	9	7	1	1
U of Chicago	Medicine	4	4	0	0
U of Illinois	Internal Med	12	10	2	0
U of Ala Birmingham	Medicine	12	10	2	0
Wash U	Ped	4	4	0	0
Baylor	Cell Bio	11	9	2	0
Childrn's Hosp Boston	Cardiol	7	7	0	0
Baylor	Col of Med	10	6	3	1
U of Wash	Pathology	22	22	0	0
Johns Hopkins	Biochem	6	6	0	0
W. Alton Jones Cell Science Ctr	Cell Science	4	3	0	1
U of Illinois	Food Chem	1	1	0	0
U of Calif San Diego	Pharmacol	12	11	1	0
Duke	Med Ctr	4	4	0	0
U of Iowa	Cardiovas Ctr	19	18	0	1
U of Calif San Fran	Metabolic Resear Unit	6	3	2	1
Childrn's Hosp Boston	Cardiology	1	1	0	0
SUNY/ Syracuse	Anatomy & Cell Bio	23	22	1	0
Boston U	Sch of Med	13	12	0	1

a Applicants were POEMB recipients
b Other Fields include: Agricultural Sciences, Health Sciences, Engineering, Psychology, Humanities and Education

Table 3.4 All POEMB Applicants: Area of Concentration for Senior Investigators with PhDs or ScDs in Biological Sciences

Area of Concentration	Number of Senior Investigators
Biochemistry	96 (46.4%)
Biophysics	11 (5.3%)
Botany	1 (0.5%)
Anatomy	4 (1.9%)
Cell Biology	4 (1.9%)
Developmental Biology/Embryology	2 (1.0%)
Biological Immunology	2 (1.0%)
Molecular Biology	16 (7.7%)
Microbiology/Bacteriology	17 (8.2%)
Microbiology	1 (0.5%)
Neuroscience	1 (0.4%)
Genetics	9 (4.4%)
Pathology, Human & Animal	3 (1.5%)
Pharmacology, Human & Animal	9 (4.4%)
Physiology, Human & Animal	15 (7.3%)
Zoology, Other	5 (2.4%)
Biological Sciences, General	7 (3.4%)
Biological Sciences, Other	4 (1.9%)
Total # of Senior Investigators with PhD or ScD in Biological Sciences	**207 (100%)**

Table 3.5 Successful POEMB Applicants: Area of Concentration for Trainees with PhDs or ScDs in Biological Sciences

Area of Concentration	Number of Trainees		
	Mass Inst of Tech Dept of Bio	U of Cincinn Dept of Microbio	U of Calif San Fran Cardiovas Resear Inst
Biochemistry	1	1	5
Biophysics	0	0	3
Anatomy	0	0	1
Cell Biology	2	1	2
Developmental Biology/Embryology	1	0	1
Biological Immunology	1	0	0
Molecular Biology	8	2	7
Microbiology	0	1	2
Neuroscience	0	0	2
Genetics, Human & Animal	1	0	3
Pathology, Human & Animal	0	0	1
Pharmacology, Human & Animal	0	1	1
Physiology, Human & Animal	0	0	1
Zoology, Other	0	0	1
Biological Sciences, General	1	0	1
Biological Sciences, Other	0	1	1
Total # of Trainees with PhD or ScD in Biological Sciences	**15**	**7**	**32**

Table 3.6 Training Awards Received by POEMB Trainees: Prior to First Year of POEMB Support

Type of Training Award	Number of Awards Received			Total
	Mass Inst of Tech Dept of Bio	U of Cincinn Dept of Microbio	U of Calif San Fran Cardiovas Resear Inst	All Three Institutes
F01 Predoctoral Fellowship	0	0	3	3
F02 Postdoctoral Fellowship	0	1	11	12
F22 Individual Research	0	0	1	1
F31 Predoctoral Individual NRSA	0	3	0	3
F32 Postdoctoral Individual NRSA	12	5	9	26
T01 Graduate Training	1	0	12	13
T05 Medical Science	0	0	1	1
T32 Institutional National Research Service Award	66	14	80	160
T34 MARC Undergrad NRSA Institutional	0	0	1	1
T35 NRSA Short Term Research Training	2	0	7	9
All Awards Above	**81**	**23**	**125**	**229**

Table 3.7 Additional Training Awards Received by POEMB Trainees: During Years of POEMB Support

Type of Training Award	Number of Awards Received			Total
	Mass Inst of Tech Dept of Bio	U of Cincinn Dept of Microbio	U of Calif San Fran Cardiovas Resear Inst	All Three Institutes
F05 International Fellowship	0	0	2	2
F31 Predoctoral Individual NRSA	0	1	0	1
F32 Postdoctoral Individual NRSA	2	2	9	13
T32 Institutional National Research Service Award	2	0	13	15
All Awards Above	**4**	**3**	**24**	**31**

Table 3.8 Additional Training Awards Received by POEMB Trainees: After the Final Year of POEMB Support

Type of Training Award	Number of Awards Received			Total
	Mass Inst of Tech Dept of Bio	U of Cincinn Dept of Microbio	U of Calif San Fran Cardiovas Resear Inst	All Three Institutes
F32 Postdoctoral Individual NRSA	6	0	2	9
T32 Institutional National Research Service Award	1	0	3	7
All Awards Above	**7**	**0**	**5**	**16**

Table 3.9 Grants or Contracts Received by POEMB Trainees: Prior to First Year of POEMB Support

Type of Grant of Contract	Number of Grants of Contracts Received			Total
	Mass Inst of Tech Dept of Bio	U of Cincinn Dept of Microbio	U of Calif San Fran Cardiovas Resear Inst	All Three Institutes
C06 Research Facilities Construction Grants	1	0	0	1
K08 Clinical Investigator Award	3	0	0	3
N01 Research and Project Contracts	0	7	10	17
P01 Research Program Projects	3	0	32	35
P50 Specialized Center	0	0	7	7
R01 Research Project Grants (Traditional)	33	0	41	74
R13 Conferences (Traditional)	1	0	1	2
S10 Biomedical Research Support Shared Instrumentation Grants	1	0	0	1
T01 Graduate Training Program	0	0	13	13
T32 Institutional National Research Service Award	0	0	27	27
TXX Research Projects Unspecified	0	0	1	1
All Grants and Contracts Above	**42**	**7**	**132**	**181**

Table 3.10 Additional Grants or Contracts Received by POEMB Trainees During Years of POEMB Support

Type of Grant or Contract	Number of Grants or Contracts Received			Total
	Mass Inst of Tech Dept of Bio	U of Cincinn Dept of Microbio	U of Calif San Fran Cardiovas Resear Inst	All Three Institutes
K11 Physician Scientist Award (Individual)	9	0	22	31
P01 Research Program Projects	0	0	10	10
R01 Research Projects Grants (Traditional)	14	0	18	32
R03 Small Research Grants	2	0	0	2
R13 Conferences (Traditional)	0	0	1	1
R35 Outstanding Investigator Grants	0	0	7	7
R37 Method to Extend Research In Time (MERIT) Award	0	0	8	8
T32 Institutional National Research Service Award	0	0	21	21
All Grants and Contracts Above	**25**	**0**	**87**	**112**

Table 3.11 Additional Research Grants or Contracts Received by POEMB Trainees After the Final Year of POEMB Support

Type of Grant or Contract	Number of Grants or Contracts Received			Total
	Mass Inst of Tech Dept of Bio	U of Cincinn Dept of Microbio	U of Calif San Fran Cardiovas Resear Inst	All Three Institutes
R01 Research Projects Grants (Traditional)	3	0	2	5
R03 Small Research Grants	0	0	2	2
R29 First Independent Research Support and Transition (FIRST) Award	3	0	0	3
All Grants and Contracts Above	**6**	**0**	**4**	**10**

4.0 Grant Application and Award Activity

4.0 Grant Application and Award Activity

In addition to expanding the base of personnel working on molecular biology applications in cardiovascular and pulmonary disease research, another major objective of the POEMB was "to foster utilization of molecular biology approaches in research areas within the mission of the National Heart, Lung, and Blood Institute where the use of these technologies has yet to be fully employed." This can be achieved in two ways: (1) investigators with an interest in cardiovascular or pulmonary disease research train in the techniques of molecular biology in order to use those techniques in their heart or lung-related studies; or (2) investigators with experience in applying molecular biology techniques to areas other than cardiovascular/pulmonary research shift their focus to heart or lung studies. To determine whether this second objective was being met, we focused on comparisons between the grant application and award activity of two major groups of researchers: individuals listed in successful POEMB applications and progress reports and individuals listed in unsuccessful POEMB applications.

The data and methodology used in this assessment are described in this chapter, followed by a discussion of the findings. Below we discuss the activity of personnel listed in POEMB grant applications (both successful and unsuccessful) in the following areas: (1) grant applications to the National Institutes of Health, (2) awards by the National Institutes of Health, (3) awards by the American Heart Association, and (4) the missionary impact of POEMB participation on those who left the program midway through. In addition, we describe the grant application activity of those researchers who left the POEMB before the end of the program in an effort to determine whether these individuals continued to pursue research that involved the application of molecular biology to the fields of cardiovascular and pulmonary disease.

4.1 Grant Applications — National Institutes of Health (NIH), Public Health Service (PHS)

This section focuses on an examination of the NIH/PHS grant application activity of individuals in the successful and the unsuccessful POEMB applicant groups and is subdivided into sections on (1) methods and (2) findings. The main goals of this examination are to describe the grant application activity of the individuals in the two POEMB applicant groups and to assess the effect (if any) of the

POEMB on the research direction/focus of these individuals.

4.1.1 Methods

The **CGAF/CRISP**[1] baseline file was used for all analyses of grant application activity reported here. Because this file contains only the names of the principal investigators (PIs) associated with each grant application and not those of support staff, all results discussed in this chapter reflect *only* the grant application activity of POEMB applicants who were listed on NIH/PHS grant applications as principal investigators. Because of the small number of matches obtained for status 5s (Support Staff) during the pilot of the NIH Index file matching, these individuals (SCODE 5) were removed from the baseline file and not considered in this analysis.

In order to make the clearest comparison of the influence of POEMB on grant awards received before and after the POEMB initiative, Battelle believed it was necessary to develop the most homogeneous measure of grant application activity possible. In this regard, we restricted analysis of research grant applications to applications for new awards (APPTYPE=1), and did not include any applications for competing or noncompeting renewal awards. The main reason for restricting analysis to new award applications is that the project period start dates (PSTART) for competing renewals are not related to the actual dates of the renewal applications but to the dates of the original application (APPTYPE=1). The variable PSTART in the **CGAF/CRISP** baseline file was used to assign applications to QUARTER. Without a valid PSTART date, competing renewal applications could not be assigned to their correct quarters.

An additional reason for excluding competing renewal applications is that these applications may be submitted after one or more successive award periods of unknown duration (but generally ranging from 3 to 5 years in length). In other words, competing renewals, although reviewed by study sections in the same ways as applications for new funding, are applications that are built upon one or more successive grant periods. These applications represent accumulations of research activity, and in this regard are essentially different from *de novo* applications. In addition, the composition of research teams may change organizationally from the initial grant award to application for competitive renewal.

---

[1] **CGAF/CRISP** (Consolidated Grant Application File/Computer Retrieval of Information on Scientific Projects).

These factors, taken together, make comparisons that combine new applications with competitive renewals difficult to interpret

Grant applications were assigned to quarters relative to each person's first POEMB application. Quarter 0 was defined so that a person's first POEMB application occurs at the midpoint of the quarter. For the purposes of this analysis, quarter 0 was included in the period before POEMB application (pre-POEMB application) and only the period from quarter -16 to quarter +16 was considered for analysis. The NIH Institute/Center/Division (ICD) code in **CGAF** was used to classify grants by NIH institution. Grants were classified as heart/lung (HL)-related if the ICD code indicated that the awarding institution was the National Heart, Lung, and Blood Institute (NHLBI) or one of its divisions. For the purposes of this analysis, all other grant applications were considered non-heart/lung (non-HL)-related. Table 2A.5 in Appendix 2A lists the seven ICD codes that were classified as HL. Also see Table 2A.6 of Appendix 2A for a list and definitions of the variables contained in the **CGAF/CRISP** baseline file. A detailed discussion of the data sets used in these analyses can be found in Chapter 2 of this report.

Battelle realizes that the timing of when researchers entered and left the POEMB program is a potential variable of interest. The only information available to us to determine this timing was the appearance of names on progress reports and the absence of those names on subsequent reports. This information was coded into a series of data source variables _1 through _7 which were used to identify investigators and trainees who left the program and determine the timing of their departure (see Section 4.4 later in this chapter). These seven variables represent possible data sources for information contained in each record of PMBPERS: _1 (= the original application, _2 (= the first progress report, _3 (= the second progress report), etc. The fields that correspond to the document(s) in which the individual is cited are tagged with an "X" as a means of tracking POEMB investigators and trainees throughout the POEMB period.

This variable, while useful for determining the timing of an investigator's departure from the program, is not reliable enough to use for characterizing all the investigators and trainees from both applicant groups as "newcomers" or "latecomers" to the POEMB program. For example, several trainees in both the successful and the unsuccessful POEMB applicant groups appear on progress reports early in the program, do not appear for several years, then reappear on one report near the end of the POEMB. Such gaps could be due to researchers taking leaves of absence or sabbaticals. Alternatively, these gaps might reflect missing information on these researchers due to an error on the part of the compilers of the

6.0 Identifying Advantages, Disadvantages, and Accomplishments of the POEMB

6.0 Identifying Advantages, Disadvantages, and Accomplishments of the POEMB

This chapter presents findings from research Battelle conducted to identify:

- Advantages and disadvantages of the POEMB approach in stimulating research activity and progress in the area of cardiovascular and pulmonary diseases.
- Significant accomplishments of the POEMB grantees programs and how the POEMB approach contributed to achieving them.

Findings are based on interviews conducted with individuals associated with the three programs supported by POEMB. A number of interviews were also conducted with individuals who organized POEMB applications for their institutions but did not receive POEMB awards. The findings reflect the perspectives of these individuals as to the value of the POEMB concept and the extent to which the three POEMB programs have been successful in achieving the main objectives established for POEMB.

This chapter is divided into three sections. Section 6.1 explains the data collection methodology used for this study. Section 6.2 presents advantages and disadvantages to the POEMB approach as viewed from the perspectives of our respondents. Section 6.3 presents accomplishments of the POEMB programs as reported by the POEMB respondents. Section 6.3 also presents findings from the interviews with the non-POEMB respondents.

Copies of data collection instruments are included as Appendix 6A, and a copy of the coding scheme used to code interview notes is attached as Appendix 6B.

6.1 Methods

We interviewed individuals among three groups of respondents: (1) principal investigators for the three POEMB programs (three respondents); (2) associate investigators in the three POEMB programs (nine respondents); and (3) principal investigators on POEMB applications that did not receive funding (eight respondents).

- **POEMB principal investigators.** This category consists of those POEMB investigators

who have program-wide administrative responsibilities, in addition to specific research responsibilities. They were typically referred to in the applications as "Program Directors." In each case, the individual interviewed was the current program director. Three POEMB principal investigators (one PI from each of the three programs) were selected for interviewing.

- **POEMB associate investigators.** This category consists of those investigators whose responsibilities were limited to research activities. These individuals are the senior-level researchers, often referred to in the grant application as "Component Leaders" or "Component Principal Investigators". Their listed responsibilities included conducting and/or oversight of research conducted under one or more of a program's components. Nine associate investigators (three from each POEMB program) were selected for interviewing.

- **Non-POEMB principal investigators.** This category consists of those individuals designated as the principal investigators on the applications submitted by institutions that did not receive the grant. Eight individuals in this category agreed to participate in this study.

Potential interviewees were identified from the Battelle-developed person-based data set created as described in Chapter 2. This database contains records for all of the investigators cited in POEMB applications, renewal applications, and progress reports available to Battelle.

In consultation with our molecular biology subject matter expert and the Delivery Order Officer, interview guides were developed for each of the three interview groups. (See Appendix 6A for copies of the interview instruments.) The interview guides were designed to address two primary types of information:

- **Advantages and disadvantages of the POEMB approach.** For all respondent groups, interviews sought the respondents' perspectives as to the advantages and disadvantages of the POEMB approach, with particular emphasis on the aspects of the program that led them to organize an application or to participate in its preparation. For the purpose of this study, we identified the five program features listed in the RFA as key aspects of the POEMB approach. (See Chapter 1, Section 1.1 for a list and descriptions of these program features.)

- **Accomplishments of the POEMB program.** The second primary type of information we sought is related to the accomplishments of the three POEMB institutions. Interviews solicited obtained the perspectives of both the principal investigators and the associate investigators on the POEMB programs as to the most significant accomplishments of their respective programs and on whether and how the POEMB concept contributed to these accomplishments. Interviews with principal investigators on applications that did not receive POEMB funding determined whether individuals in this group of respondents were aware of any important accomplishments by the programs that received the award, and further, whether they or any of their colleagues had maintained an interest in the application of the techniques of molecular biology in the areas of cardiovascular and/or pulmonary diseases.

Open-ended, semi-structured interviews were conducted over the telephone. Documentation of each interview consisted of both hand-written notes of what was discussed and an audio tape recording to ensure that complete records of all the interviews were available for referral.

For each completed interview, a transcript was developed based on the notes taken during the interview and the audio recording. The transcript was then imported into the qualitative data analysis software package, The Ethnograph®, as a data file. Using The Ethnograph®, a content analysis of the interview data files was conducted. (See Appendix 6B for the code book and a summary of the coding scheme used in the analysis of the interview contents.) By comparing the contents of interview data files, similarities and differences within and across the interview groups were identified.

6.2 Advantages and Disadvantages of the POEMB Approach

In this section we explore perceived advantages and disadvantages of the POEMB approach as viewed from the perspective of three different categories of respondents: (1) POEMB principal investigators, (2) POEMB associate investigators, and (3) non-POEMB principal investigators. Sections 6.2.1 and 6.2.2 deal with perceived advantages, and Section 6.2.3 with perceived disadvantages.

6.2.1 Advantages of the POEMB Approach

One of the primary objectives of this study was to obtain the perspectives of our respondents (n=20) as to the value and advantages of the POEMB approach. In particular, we sought to understand

what specific aspects of the POEMB approach respondents considered valuable or advantageous and what led them to either organize an application or participate in its preparation.

In general, the main motivation to apply for the POEMB grant was the desire to pursue what was seen as a new and important direction in heart and lung research, specifically, to apply the techniques of molecular biology to the cardiovascular and pulmonary systems in ways that had not been done before. How the new direction was going to be pursued depended on the focus of research and the expertise of the institutions' researchers at the time the decision to submit an application was made. For example, the principal investigator and associate investigators at one POEMB institution explained that the grant presented the opportunity to bring all the molecular biologists in the department together under one program to pursue cardiovascular research, a new area for many of the faculty. The investigators at another POEMB institution wanted to bring experts in molecular biology together with faculty from an established center for cardiovascular and pulmonary research.

The principal investigators with the non-POEMB institutions organized their applications because they saw an opportunity to develop further what were already established cardiovascular and pulmonary research centers. There was the perception among our respondents in this category that molecular biology was becoming increasingly important to cardiovascular and pulmonary research and that the POEMB would be a good vehicle for promoting the long-term development of their institutions in this direction.

6.2.2 Relative Importance of Key Features of the POEMB Approach

In addition to the more general discussions of the value of the POEMB approach, respondents discussed specific aspects of the POEMB approach they considered valuable or advantageous, and whether or not these aspects had prompted them to organize or participate in the submission of an application. Each of the five key features are discussed below in terms of respondents' perceptions of them. Table 6.1 tabulates the number of respondents (by respondent type) who indicated they consider a given feature important.

Table 6.1 Respondents Who Considered a Given Feature of the POEMB Important

Feature	POEMB PIs (n=3)	NonPOEMB PIs (n=8)	POEMB Investigators (n=9)	Totals
Seven-year award	3	8	9	20
Recruitment of essential expertise	2	8	6	16
Institutional environment and resources	3	6	7	16
Research freedom	(2*)	1	1 (5*)	2 (7*)
New investigators	3	6	6	15

* Indicated feature was important only in connection with other features.

6.2.2.1 The Seven-Year Award

All respondents reported that the seven-year award was an attractive feature, although for three of the POEMB associate investigators and three of the non-POEMB PIs this feature was of secondary importance in relation to other features. For those respondents who considered this feature to have been of primary importance, an extended (and extendable) seven-year award period was considered crucial to pursuing long-term goals that would be more difficult to achieve with conventional forms of support. This was seen as an advantage in that it would allow researchers to attempt innovative types of research that require a long-term investment of time and resources in order to see results. Moreover, the seven-year award provided them the security of a consistent funding source for at least seven years.

Several respondents linked the seven-year award to other features, emphasizing the fact that this feature became even more valuable in relation to other features. Five associate investigators reported that the seven-year award and the broad degree of research freedom were inextricably linked. As they saw it, the ability to try innovative approaches and explore new areas would be dependent not only on having the skills and resources, but also on having the time to apply them. Three of the non-POEMB investigators anticipated that the longer than usual award period would complement the support for new investigators, providing a more secure transition period for post-doctorate PhDs.

6.2.2.2 Support for Recruiting Essential Scientific Expertise

Nearly all of the respondents (16) reported that this feature was important in their decisions to apply for the POEMB grant. Two POEMB PIs and six associate investigators considered this feature valuable, and all of the non-POEMB PIs concurred. However, these respondents were generally more interested in bringing together groups of molecular biologists and cardiovascular/ pulmonary specialists from within their respective institutions, rather than recruiting new faculty from other institutions. In other words, there was a perception that POEMB would facilitate the creation of a more interactive and collaborative environment where researchers from independent laboratories would be able to share resources and talents. Only one POEMB PI and three associate investigators reported that this feature was not important to their decisions to be involved in their respective applications.

6.2.2.3 Support for Institutional Environment and Resources

Sixteen respondents reported that the opportunity to enhance their institutions' research facilities and other resources was a very attractive feature at the time they were organizing their applications: three POEMB PIs, seven POEMB associate investigators, and six non-POEMB investigators.

Respondents had different needs regarding technical facilities at the time of the application, and this influenced how they perceived the value of this feature of the POEMB grant. The application of molecular biology techniques requires facilities that are expensive and complex (e.g., a transgenic facility for the generation of transgenic knockout animals; a core cell biology facility for the analysis of the knockout animals), and it is rare for an institution to receive financial support that can be dedicated to the development of such facilities. Therefore, those who did not have access to these facilities prior to the POEMB application considered POEMB unique in that it explicitly allowed institutions to devote a significant proportion of grant funds to purchasing new equipment and overall development of facilities. Two of the POEMB associate investigators we interviewed already had the necessary equipment in their laboratories prior to the application, so this feature was not especially important to them. Two of the non-POEMB investigators did not mention this feature at all.

6.2.2.4 Broad Degree of Research Freedom

The perceived value of this feature varied among respondents, with only one associate investigator and one non-POEMB investigator reporting it as an important feature. Two POEMB PIs and five associate investigators discussed the value of this feature, but only in relation to other features of the grant, e.g., support for new investigators, cross-disciplinary collaboration, and the seven-year award.

This feature was not reported as important by nine of the respondents. One POEMB PI simply did not mention this feature in his discussion of the aspects of the POEMB that he considered advantageous. A POEMB associate investigator dismissed this feature entirely, asserting that there was insufficient research freedom and too much interference in his work. Seven non-POEMB investigators and one POEMB associate investigator reported that the feature was not particularly important because it was not unique to POEMB, and thus was not an aspect that attracted their attention and motivated their response to the RFA.

6.2.2.5 Support for New Investigators

Most respondents reported that they viewed the support for new investigators as a valuable feature of the POEMB approach and that it was an important factor in their decision to organize or participate in a POEMB application. A total of 15 respondents described this feature as valuable: all three POEMB principal investigators, six POEMB associate investigators, and six non-POEMB investigators. Only five respondents reported that they did not consider this feature important when they decided to join their institutions' application efforts. Among the reasons respondents cited for considering this feature valuable are the following:

- *Emphasis on training within a collaborative and innovative research environment.* Respondents from both POEMB and non-POEMB institutions shared NHLBI's goal of providing opportunities for new investigators to become skilled in the approaches of molecular biology and to apply these techniques to cardiovascular and pulmonary disease research. While training in this area can occur in many different contexts, POEMB was

considered unique because it combined training with the creation of a collaborative research environment where innovative approaches to heart and lung research would be pursued.

- *Emphasis on the training of MDs in basic science.* Two PIs and one associate investigator from POEMB institutions reported that because of their backgrounds in clinical work they had a special interest in training MDs in basic science, particularly in the application of molecular biology to heart and lung disease research.

- *Alternative model to the traditional two-year post-doctorate.* Three POEMB associate investigators and two non-POEMB investigators reported that they saw a need for an alternative form of support for post-doctorate PhDs. The perception was that the traditional two-year post-doctorate does not provide enough time or security to PhDs who are in the transition period to an academic faculty position. The POEMB provided new PhDs with a longer, more secure period during which they could develop their own lines of research, learn new skills, and eventually become more competitive in terms of acquiring their own funding and obtaining an academic job.

6.2.3 Disadvantages of the POEMB Approach

Another objective of this study was to obtain the respondents' perspectives as to the disadvantages of the POEMB approach at the time they were involved in organizing or participating in their institutions' application efforts. Eleven respondents reported that they did not perceive or anticipate any disadvantages from the POEMB approach: two POEMB principal investigators, six POEMB associate investigators, and three non-POEMB investigators.

Meanwhile, nine respondents reported that they did perceive or anticipate disadvantages with the POEMB approach: one POEMB principal investigator, three associate investigators, and five non-POEMB investigators. The *potential* disadvantages were related to a number of different concerns. These are:

- *High degree of sponsor control over research.* The POEMB PI and a non-POEMB investigator both expressed concern that NHLBI would assert a high degree of control over the research that would be conducted under the auspices of the grant. However, the POEMB PI explained that this never actually materialized as a problem.

- *Establishing and maintaining collaboration.* Three POEMB associate investigators and one non-POEMB investigator foresaw problems with establishing and sustaining successful

collaboration within their proposed research groups. One of these associate investigators explained that, despite his concerns the collaboration worked out very well in his program, while the other two (both at the same institution) stated that collaboration within their program did prove to be difficult to develop and sustain.

- *Inadequate funding levels.* One non-POEMB investigator thought that the funding levels would not be adequate for what was expected of the programs.

- *Administrative difficulties with the application and review process.* Another non-POEMB investigator discussed disadvantages in terms of the application and review process.

- *New investigators less competitive than their counterparts on R01s.* A third non-POEMB investigator speculated that the new investigators trained through POEMB would not be as competitive careerwise as those who had led their own R01s rather than joining a program project such as POEMB. This concern contrasted with what the majority of POEMB investigators described as a principal advantage of the grant, namely that it would support new investigators during the transition from fellowships to independent research and would actually increase their competitiveness.

6.3 POEMB Program Accomplishments

This section of Chapter 6 is divided into five main subsections. The first subsection examines — from the perspective of POEMB participants — the success of the POEMB program in meeting its primary objectives. Also from the participants' perspective, the second section explores how key features of the POEMB grant program have contributed to that success. The third explores instances of collaboration—both intra- and inter-institutional — stimulated by the POEMB program. The fourth examines whether the principal investigators at the institutions that applied for but did not receive the grant are aware of any important accomplishments on the part of the POEMB institutions, and whether they or any of their colleagues have maintained an interest in the application of the techniques of molecular biology in the areas of cardiovascular and/or pulmonary diseases. A final summary addresses the overall impact of the POEMB program.

6.3.1 Accomplishments of the POEMB

We asked respondents to report accomplishments of their respective POEMB programs toward achievement of the two primary objectives: (1) to foster utilization of molecular biology approaches in cardiovascular and pulmonary research and (2) to provide opportunities for new investigators to become skilled in the application of such approaches to cardiovascular and pulmonary research.

6.3.1.1 Utilization of Molecular Biology Approaches in Cardiovascular and Pulmonary Research

Principal investigators among our respondents provided overviews of their institutions' accomplishments, identifying as particularly significant the following scientific achievements by researchers in their institutions' Programs of Excellence:

- Work on lung surfactant proteins aiding in premature neonate therapy, and the generation of a working cystic fibrosis (knockout) mouse model applicable to human disease and eventual gene therapy.
- The cloning of 30 transport genes and the study of their function and effects on blood pressure control.
- The characterization of changing levels of myosin production and their effects on cardiac tissue function.
- The generation of adrenergic receptor mutants to help define their role in asthma.
- The isolation of cardiac homeobox genes and generation of knockout derivatives.
- The generation of knockout derivatives of the growth factor TGF-β1 gene and its cardiomyopathic consequences.
- The identification of a variant of sodium ion channels, specific for African-Americans, which play a role in hypertension.
- The first identification and definition of promoters and other genetic regulatory regions involved in arterial wall and platelet structure and function, a tool that may assist in the examination of endothelial cell diversity and in the development of animal models of thrombosis (the complex series of cellular and molecular events that lead to blood clotting), and the development of antisense oligonucleotides to treat arterial disorders.

- The isolation of a long-chain fatty acid transport protein (FATP) involved in energy production in cardiac cells, and the description of receptor trafficking mechanisms that may be germane to other cardiovascular signaling mechanisms.

- The definition of macrophage scavenger receptors that appear to play a role in the deposition of lipoprotein and cholesterol in arterial walls, and the first molecular cloning of a well-defined HDL (high density lipoprotein) receptor which may play an important role in cholesterol transport.

- The development of connective tissue disease animal models, and the development of a new model describing the pathway for collagen metabolism in blood vessel walls.

- The demonstration of adhesive proteins and receptors, and their role in the development of heart blood vessels, and the generation of concrete evidence for the role of adhesive proteins in the recruitment of inflammatory cells in plaque formation.

- The first description of myotonic dystrophy, a cardiac genetic disorder.

- The development of molecular tools for somatic gene therapy, the first "knockout" mice in the fibrolytic system establishing its role in the development of arterial and venous thrombuli (blood clots).

- X-ray crystallographic work performed on the three-dimensional structure of cardiac proteins (e.g., Kinesin, thyroid hormone receptor, and keratinocyte growth factor).

- Embryonic mesoderm induction (muscle gene switching during differentiation) found to be extrinsically controlled. Investigators monitored the early contractile protein gene, cardiac tropnin T.

- Work done by a *Drosophila* geneticist on embryonic cell cycles found to be relevant to the study of the molecular biology of cardiac cell cycles.

Associate investigators among our respondents covered a subset of the POEMB program accomplishments in greater depth. These are reported below.

The development of a set of reagents (nucleic acid probes) for the study of various genes in the pulmonary myocardium. (ANF, MLC1a, MLC2v, α-MyHC–cardiac tissue genes that play a part in heart and lung development). These reagents have been made available to over 400 investigators in the US and internationally. The transcriptional control of the above genetic elements, and their knockout derivatives, can now be used to perform structure/function relationship experiments pertaining to their

role in heart development.

Animal models using ES (embryonic stem) cell gene targeting technology. At the time of the initial RFA there were no animals made with this technology. The new animals are becoming important in determining the physiological function of the proteins under study (sodium and hydrogen ion exchangers); results that probably would have been otherwise unobtainable. A specific example of this was the cloning and generation of a knockout mouse harboring a mutated calcium pump gene that affected aortic vascular muscle contractility, a phenomenon not predictable other than with an appropriate animal model.

The characterizing of promoter subsets of genes (lung surfactant proteins and epithelial tissue) describing lung cell functions in transgenic animals. This has led to a better understanding of the role of these genes in development and disease. Also stated was the development of promoter elements allowing gene ablation and addition experiments. A start was made in the cloning and characterization of a series of transcription regulator genes that control lung development. They turn out to be important not only in lung organogenesis, but also in cystic fibrosis. Lastly, a series of lung epithelial surfactant homeostasis genes (*spA-D*) were cloned or characterized. Targeting of these genes is in progress and the anticipated models are expected to provide insights into the role of each of these proteins in lung function. A quote from one of the associate investigators speaks to the importance of the POEMB in making this particular accomplishment possible, "Having a user-friendly program has made it wonderful to take on the gene targeting program without fear of flopping."

The cloning of scavenger receptors, which led to the cloning and characterization of the first HDL receptor. Also genes delineating the function of the Golgi apparatus were cloned. The investigator who reported this accomplishment stated that the cloning and characterization of the HDL receptor would not have occurred without the POEMB program. Some of the contributing aspects of the POEMB program were the availability of trainees, the enhanced facilities, and the stability afforded by the seven-year award.

The analysis of a great number of animal models (specifically, mice) for vascular biology were performed. Additionally, the large number of published mouse strains ("knockouts") and other immunological reagents were made available through commercial enterprises, e.g., Jackson Labs. Clones and antibodies were made available after publication. Most of these accomplishments were dependent on the new facilities created through the POEMB.

The cloning of the unique fatty acid transport protein (FATP) expressed in cardiac muscle and fat cells. This work was originally proposed as a search for glucose transport proteins, important in energy utilization. Prior to this study, it was not known how fatty acids were transported into cells. Thus this new clone may represent an entirely new class of transport proteins. The work was performed over a two-year period by a POEMB trainee — an MD who has specialized in cardiology. An article on the work was published in the journal *Cell*, an internationally respected, peer-reviewed journal of eukaryotic molecular biology.

POEMB is credited with providing the resources necessary for developing the cloning strategy, i.e., "expression cloning," a technique whereby a library of clones is screened in a series of cell lines, searching for, in this case, an enhanced fatty acid uptake. The FATP clone expressed the desired phenotype. POEMB provided for the facilities needed in these experiments and support for the trainee who actually conducted the bulk of the work. This in turn contributed to her career development as a researcher skilled in molecular biological techniques.

Attempts were made to resolve the structure of proteins involved in the developmental biology of the cardiovascular system. A large number of proteins were involved and not all of the projects were successful. One success was the lung surfactant protein SP-A carbohydrate binding domain. A second (hypoxygenases) was not successfully completed within the POEMB program, but was eventually transferred to a commercial concern that is now involved in pharmaceutical inhibitors of the enzyme. Two other proteins were structurally characterized, and these were also supported by ancillary funding. These were the thyroid receptors from rat and human, which contributed to an understanding of how nuclear receptors function in the development and physiological function of cardiac cells. The other was the structural determination of a class of kinesin motor proteins. These

proteins are involved, among other functions, in separating chromosomes and in delivering proteins to mitochondria.

Basic cell cycle biology of *Drosophila* was applied to topics that may be relevant to cardiovascular and pulmonary research. The research focused on the appearance of cell division stage G1 (the first quiescent stage of cell division after the active "S" stage). The research has elucidated a diversity of cell division control elements active at specific cell cycle stages. Genes identified in such control included cyclin E, the transcription factor E2F, and a repressor of the transcription factor RB (itself a repressor of replication). All of these genes are apparently conserved in mammals and *Drosophila*, "and constitute important access to cell cycle control" [in mammals] particularly with respect to oncogene expression. The work showed that the transcription factor, E2F, is activated independently of, and inhibited by, cyclin E, and that cyclin E is also required for the downstream activation of the transcription factor. Thus a loop instrumental in the regulation of DNA replication, and ultimately cell division, has been defined.

Development of mammalian pulmonary epithelial tissue; specifically with the distal pulmonary epithelium comprised of Type 1 cells which are involved in gas exchange, and Type 2 cells which produce surfactant proteins. During fetal lung development, certain populations of Type 2 cells differentiate ("transdifferentiate") into Type 1 cells. This investigator's group used the promoter of a glycoprotein cell surface marker (RT1) to monitor this transition. Also, there was work on the effects of glucocorticoids (a class of hormones) on the levels of mRNA of four surfactant genes. Each gene apparently had a unique response to the same stimulus; thus implicating the promoter regions of each gene in the separate and differential responses. The associate investigator explained that they were not able to address all of the component's goals in, what was his opinion, a fully satisfactory way. Nevertheless, they were able to define new and important directions for lung epithelial tissue transdifferentiation research.

6.3.1.2 Support for the Training of New Investigators

Principal investigators among our respondents reported considerable success in the training of new investigators. In the first seven years of POEMB at one institution, for example, 54 postdoctoral fellows were placed in first-rate academic institutions: 12 MD/PhDs, 18 MDs, and 24 PhDs. Thirty percent of the new investigators trained at this institution are women. Several of the early trainees have already achieved full professor status at various institutions.

Associate investigators stressed similar accomplishments in the area of training new investigators.

- One respondent in this category pointed out that many of the POEMB trainees from their institution have gone on to productive careers in cardiovascular and pulmonary research and molecular biology:
  - An MD/cardiologist has gone on to become an assistant professor of medicine at a major university, specializing in cardiology and cell biology.
  - Another PhD trainee, who did work on the HDL receptor, is now an assistant professor of a large urban medical school.
  - Still another former trainee is a full professor in Japan, a remarkable accomplishment in and of itself.
- Another associate investigator reported that the bulk of the work for one of their components was carried out by one of the junior investigators within the POEMB. In this capacity, the junior investigator was allowed to develop a relatively independent program of research in which he could study the differentiation of the development pathways of lung cells. His work in this area was described as having a positive impact on his development as a researcher.
- A third respondent in this category described a monthly series of seminars on structural biology research that brought together junior investigators with different backgrounds and interests, facilitating familiarity and comfort with a multidisciplinary setting among the trainees.
- Another of the POEMB institutions used yearly workshops as a vehicle for the dissemination of molecular biological techniques, an activity involving reciprocal exchanges between basic molecular biologists and MDs. Some of the early trainees are now featured speakers.

In the context of this objective, "new investigators" need not always be junior-level investigators, but can also be investigators new to a particular research area. One associate investigator offered examples of researchers who had crossed disciplines through their work on POEMB projects.

- Several individuals with molecular biological expertise in other disciplines have been trained and now focus on cardiovascular and pulmonary research. One yeast geneticist is now working in the cardiovascular system.
- Another *Drosophila* geneticist working on *hox* genes has found analogous genes important in early cardiovascular development.
- A third individual previously working on retinoids (compounds of the carotenoid and vitamin A family) has employed some of the promoters discovered in this component to express retinoids in the heart and has potentially elucidated therapeutic value for these compounds (i.e., the closure of ductus arteriosi). Additionally, the fruits of the labor, both in terms of the research results and the reagents, have been made readily available to the scientific community.

6.3.2 Contribution of Key Features of the POEMB Approach to Accomplishments

We asked respondents to discuss how the POEMB approach contributed to their programs' accomplishments. More specifically, we asked respondents how the key features of the POEMB approach supported the achievement of program objectives.

6.3.2.1 Seven-Year Award

Respondents in both the principal investigator and the associate investigator categories felt the seven-year award was a critical feature of the POEMB program. Especially important during the early days of the grant when investigators were exploring potential avenues of research, the longer award period allowed researchers to pursue more complicated projects that could not be expected to be as fruitful in the short term. Associate investigators also acknowledge that the seven-year award played a role for them, although the length of the award seemed less critical for some of them than for the principal investigators.

6.3.2.2 Support for Recruiting Essential Scientific Expertise

In their discussions of this feature, POEMB investigators almost invariably viewed the grant as a means to increase interaction and collaboration among existing faculty, rather than as a means for the recruitment of new, additional faculty. Respondents reported success at bringing into their programs both established cardiovascular and pulmonary researchers and scientists already employing molecular biological techniques in other departments. A POEMB PI offered an example of a major career change made by one POEMB researcher:

- One of the associate investigators, who was previously involved in molecular biology research on skeletal muscle, was brought in to the program to do research on heart and lung tissue. He is now the director of cardiovascular biology at a children's hospital.

In general, the associate investigators tended to stress the importance of the POEMB-stimulated interaction and collaboration to a greater extent than did their PIs. Respondents explained that the ability to continuously engage their colleagues enriched their own understanding of applying molecular biology in cardiovascular research by allowing them to learn from their respective successes and failures.

- One of the associate investigators mentioned in particular his collaboration with an expert in transgenic technology.
- A second emphasized the increased cooperation and interaction among the researchers that synergized the overall work done at the institution.

Intra-institutional POEMB meetings and seminars were frequently a focus for sharing information and techniques among the different labs.

6.3.2.3 Support for Institutional Environment and Resources

Principal investigators among our respondents attributed the development of important core facilities to support received from POEMB. At one institution, for example, the new facilities whose existence were attributed to POEMB comprised:

- A DNA Core for nucleic acid synthesis, sequencing, and analysis.
- A Transgenic Animal Core used for the construction of new genetic animal (mouse) strains and knockout mutants.
- A Cell Biology Core for the analysis of new genetic constructions (e.g., the knockout mice).

Researchers at the POEMB institutions but not formally associated with the program have benefitted as well from both the facilities and the new genetic animal strains.

Associate investigators pointed out that access to these POEMB-supported facilities gave researchers the freedom to pursue new research questions in a much more rapid fashion than would otherwise have been possible. Moreover, the new facilities provided the perfect training ground for teaching molecular biology techniques to new investigators.

Two POEMB PI respondents mentioned the importance of a well-developed research infrastructure to an institution's ability to attract future funding from both government and commercial sources, bringing their institutions increased recognition and credibility in the POEMB-supported fields.

- New cardiovascular facilities were constructed at one POEMB institution and ancillary funding was obtained from the state and from a private foundation ($8 to 9 million). The result was a major increase in the research devoted to cardiovascular and pulmonary research.
- Another respondent noted that several new R01 projects have already developed out of the program.

6.3.2.4 Broad Degree of Research Freedom

Principal investigators among our respondents appreciated the broad degree of research freedom allowed by the POEMB, but chiefly in combination with other key features, such as the long award period and the facilities enhancement support. Associate investigators generally concurred, although several at one institution either felt the broad degree of freedom was non-existent, dismissed it as a structural flaw, or simply did not mention it at all. One respondent in particular did not see an increased

amount of research freedom associated with POEMB, stating that he prefers what he sees as the freedom of R01-type grants. Another of the associate investigators at this institution saw the broad degree of research freedom offered by POEMB as a structural problem. He felt that the inherent lack of specificity and focus in the POEMB objectives in turn led to an incoherent institutional group, which itself lacked focus and dedication to a set of specific, shared goals.

6.3.2.5 Support for New Investigators

Efforts made by the POEMBs to provide training for new investigators was an aspect of their programs of Excellence of which principal investigators seemed especially proud. POEMB was seen as an important recruitment vehicle, attracting talented young researchers to the POEMB programs, and respondents felt that the new investigators both benefitted from the POEMB and contributed to its accomplishments. One PI pointed out that trainees at his institution who had become skilled in the techniques taught in the core facility labs had subsequently taken this knowledge to other institutions and set up their own transgenic facilities, with what he termed "an incredible multiplying effect on taking this technology and distributing it around the country."

The recruitment and training of post-graduate MDs was an issue that received mixed reviews. One PI favored NHLBI's encouragement in this direction and discussed the importance of having "MDs working side-by-side with PhDs." He felt this cross-disciplinary staffing has enriched all the program's labs at his institution because of the combination of basic research expertise offered by the PhDs and the clinical knowledge brought in by the MDs. According to the PI, this has proved particularly helpful in the analysis of pathology in the animal models generated by the transgenic facilities. The other PIs agreed with the basic principle of recruiting MDs, but found it difficult to recruit them in the numbers they felt were expected by NHLBI.

One of the associate investigators among our respondents strongly disagreed with POEMB's programmatic emphasis on recruiting and training MDs. As he understood it, the premise behind this component of the program was that in order for research to be medically relevant, MDs should be trained in basic research, a premise he fundamentally disagreed with. In his opinion, people who are trained in basic research (i.e., PhDs) should be the ones studying medically relevant problems.

6.3.3 Collaboration Facilitated by POEMB

Although promoting the development of collaborative efforts was not one of the POEMB's primary objectives or key programmatic features, it was clearly of importance to the program's designers and was specifically mentioned as a programmatic goal in the RFA. Most respondents felt that collaboration works best intra-institutionally. Many were far less certain of the usefulness of the annual meetings set up by NHLBI to foster inter-institutional interactions between the investigators of the three POEMB institutions. The competitive nature of scientific endeavor may be one barrier to such inter-institutional cooperation. Other possible barriers noted by our respondents include physical distance between institutions, lack of commonality of interest, and resistance to "forced" interaction.

Despite their rather negative reactions as a group to the formal annual POEMB meetings held to facilitate inter-POEMB relationships, individual associate investigators did offer a number of specific examples of cross-site collaborations that had sprung up.

- One investigator shared cDNA clones with a lab from another POEMB institution doing work on plasmembrane calcium pump expression during smooth muscle cell cycle. He also offered to share knockout mice for the plasmembrane calcium pumps, as soon as those were developed, for use in studies of smooth muscle.

- Another investigator "coordinated publications" with an investigator in one of the other POEMB programs.

- A third investigator mentioned five or six publications their group has co-authored with different POEMB groups, several with each of the other two POEMBs. Some of these collaborations are still active and ongoing.

- Examples of cooperation cited by a fourth investigator included shared reagents and protocols. He and his colleagues taught one of the other POEMB programs how to do *in situ*'s, and that program, in turn, helped them with some of the cytokine modeling, some of the assays. Clones were also exchanged.

- A fifth investigator mentioned a surfactant protein project that was "put on track" by materials provided by one of the other POEMB programs. This investigator felt that if success were experienced by year's end, it would be because of this interaction.

- A group at one of the POEMB programs did not have much physiologic experience with

the animals. They were able to send their post-doc to a lab at another POEMB program for three days of learning and interaction in this area.

The POEMB PIs concurred that informal networking among individual researchers was more fruitful than the formal group meetings:

- One pointed to specific support his program had received from the other two POEMB programs, in terms of sharing of vectors and promoters.
- A second felt that the collaboration NIH was seeking might have been developing among the younger researchers, but that its effects would only be visible in the long term.

6.4 Impact of the POEMB Beyond Program Participants

In an effort to explore the impact of the POEMB program beyond its effect on program participants, we interviewed eight principal investigators whose institutions had applied for but not received POEMB grants. A number of these principal investigators as well as their fellow investigators have maintained an interest in the application of the techniques of molecular biology in the areas of cardiovascular and/or pulmonary biology. The majority were also aware of important accomplishments by the programs that received the award. However, most respondents in this category tended to feel that (1) the accomplishments of POEMB program personnel related more to the caliber of individual POEMB investigators than to their participation in POEMB and (2) similar achievements could also be pointed to among non-POEMB investigators.

Interestingly, three of the investigators from non-POEMB institutions who have maintained an interest in this area remarked that the POEMB application effort itself actually served as an impetus for their increased interest in multidisciplinary approaches to cardiovascular and pulmonary biology.

- One non-POEMB investigator noted that "even going for the funding does help to mobilize people within the institution and put them together." He later explained that his own cardiovascular center has experienced an accelerated evolution into cell biology and molecular biology in the last five or six years (although a direct relationship between the application effort and this evolution was not explicitly stated).
- Another of the non-POEMB investigators explained that the application effort had a long-

lasting effect at the institutions that sponsored the application she organized. In this case, there were two results:

— The university arranged for new facilities for the department, moving it into a new building.

— A number of people went into cardiovascular or lung research who had previously not considered doing so.

One respondent from a non-POEMB institution asserted that the three POEMB programs have not been productive in the field of cardiovascular and pulmonary research. Although the expertise and reputation of the POEMB investigators is acknowledged, he does not think significant work has come from any of them. He believes that the emphasis on molecular biology at these institutions has been misdirected, and that one should not look to solve medical problems simply because of an intense, general interest in molecular biology, as he claims has been done at the three programs. It is important to note that this respondent had maintained only a partial interest in the application of molecular biology to cardiovascular diseases and that the approach is not a major part of his own research.

6.5 Overall Impact of the POEMB Initiative

Associate investigators among our respondents at one of the institutions tended to disagree with the way certain aspects of the POEMB were administered. Yet even the most outspoken of these individuals felt that his own component was generally successful and that the structure of the POEMB did not hinder his research in any significant way.

Most POEMB respondents (both PIs and associate investigators), when asked whether the accomplishments could have occurred without POEMB, reported that, perhaps given the right combination of funding and support, similar achievements might have been possible. Most felt, however, that this same level of support would have been very difficult to assemble.

Descriptions such as those below closely approximate the outcome envisioned by designers of the POEMB program:

- One PI stated that POEMB had contributed to his institution by generating what he termed a "community effect." The collaborations and interactions that occurred between

university researchers enhanced the general level of knowledge of cardiovascular biology and how molecular biology could be applied to this area. The investigators of his institution's POEMB were able to participate in symposia with researchers from other regional universities on these issues.

- The principal investigator at another institution characterized the POEMB grant and the research groups it spawned as a nucleus for scientific talent, infrastructure and facilities support, and other funding that enabled molecular biological approaches to cardiovascular and pulmonary research that otherwise would not have occurred.

Yet perhaps, as several of our respondents suggested, the full impact of the POEMB program will be seen only in future years as the cadre of new investigators trained by the program begin to establish laboratories of their own in which to train others.

References Cited

References Cited

Bickel, PJ, Hammel, EA, & O'Connell, JW. Sex Bias in Graduate Admissions: Data from Berkeley. *Science* 1975; 187:398-404.

Cochran, WG, & Cox, GM. *Experimental Designs*. New York: John Wiley & Sons, Inc., 1975.

Littell, RC, Milliken, GA, Stroup, WW, & Wolfinger, RD. *SAS® System for Mixed Models*. Cary, North Carolina: SAS Institute, Inc., 1996..

Medical Subject Headings (MeSH) *Annotated Alphabetical List*. Bethesda, MD: National Library of Medicine, 1997, pp. I107-122.

Morgan, HE, & Paul, SR. American Heart Association-Bugher Foundation Centers for Molecular Biology in the Cardiovascular System. *Circulation* 1995; 91:487-493.

Office of Science Policy and Technology Transfer. *Documentation 1994 Consolidated Grant Application File, Archive File*. Bethesda, Maryland: National Institutes of Health, 1994.

Ulrich's International Periodicals Directory. Medical Sciences- Respiratory Diseases. New Providence, NJ: RR Bowker, 1996, pp. 4592-4593.

CRISP Intramural Research Index. Washington, DC: United States Department of Health and Human Services, 1989.

Appendix 2A

Data Sets Documentation

Appendix 2A

Data Sets Documentation

List of Tables

Appendix 2A

Structure of the Battelle-Developed Person-Based Database

Following is a summary of the relational database (**PMBPERSON**) designed by Battelle to house the data obtained from POEMB applications and progress reports on the staff being proposed for or actually working on the POEMB grants. This Battelle-developed relational database, which was used for matching against larger NIH and other related data sets of interest to this study, consists of six linked tables (see Table 1).

Table 1
Components of the Battelle-Developed Person-Based Database

Name of Table	Description
PMBAPP	POEMB applications received by NHLBI.
PMBSITE	Institutions named as performance sites in POEMB applications.
PMBPERS	Scientific staff cited in POEMB applications and progress reports.
PMBALT	Alternative data on individuals with records in **PMBPERS**, as noted by Battelle staff during data entry, cleaning, and verification procedures.
PMBINALT	Alternative data on individuals with records in **PMBPERS**, as reported by the applicant institutions during institutional verification procedures.
PMBED	Educational experience of staff listed in POEMB applications and progress reports.

The following fields were used to link the tables: ACODE and ICODE (contained in **PMBSITE, PMBAPP,** and **PMBPERS**) and PCODE (contained in **PMBPERS, PMBALT, PMBINALT,** and **PMBED**). Figure 2.1 depicts the links between each of these six databases.

PMBAPP

PMBAPP is a table for Round 1 and Round 2 applications received by NHLBI in response to the POEMB Request for Applications (RFA). Data entry into this table was completed on the 18 applications (Phase 1 applications) provided to Battelle during our first meeting with NHLBI staff, as well as on eight applications received subsequently (Phase 2 applications). **PMBAPP** currently contains 26 records, corresponding to the 26 applications Battelle received. The actual number of non-duplicated applications was 24, which means that of 27 POEMB applications submitted to NHLBI, all but three were forwarded to Battelle for inclusion in this study.

There are two duplicates among the 26 applications in **PMBAPP**. After data entry had been completed, it was discovered that Applications 14 and 19 were exact duplicates, as were Applications 6 and 20. However, names (i.e., records) duplicated in the paired applications were subsequently culled from **PMBPERS** before matching with the larger data sets began.

This table is linked to **PMBSITE** (the table of POEMB performance sites) and **PMBPERS** (the table of individuals cited in POEMB applications and progress reports) through the fields ACODE and ICODE.

Fields in the **PMBAPP** table include ACODE, ICODE, APP#, SS#, APPRCVD, and REVSCORE. These fields are described below.

ACODE.

Upon receipt of the POEMB applications from NHLBI, Battelle assigned to each an identifying number or ACODE (application code). ACODES 1-3 were assigned to the three winning proposals.

		APP #	ACODE
Rosenberg, R	MIT	41484	1
Lingrel, J	U Cincinnati	41496	2
Williams, L	UCSF	43821	3

Other codes were assigned to the remaining applications in no particular order. No attempt was made to assign ACODES in chronological (or any other) order. The field APPRCVD nevertheless allows us to distinguish between Round 1 and Round 2 applications from the same institution. When two applications were received from a given institution (i.e., Round 1 and Round 2), each application was assigned a different number.

In several cases, NHLBI received Round 1 and Round 2 applications from a single institution. The pairs of applications tended to vary considerably — different PIs, different key personnel, different performance sites, and different proposed research components. We therefore treated them as distinct applications, although duplicate names were purged from the person-based file before matching against the larger data sets was attempted.

ICODE.

Battelle composed an ICODE (institutional code) by combining the ACODE and the SITE# fields in the **PMBSITE** table. The ICODE that appears in **PMBAPP** is the code that identifies the *principal site*, listed on the cover page of the POEMB application under "mailing address." For this reason, all ICODES in **PMBAPP** end in zero (0). (See discussion of ICODE under **PMBSITE** below.)

APP#.

Data for the field APP# is taken from the number assigned by NHLBI to each application (available from the cover page on page 1 of the application). These numbers always begin with the characters "1 P01 HL," which were data-entered automatically. This is an important field for interfacing with the larger NIH data sets.

SS#.

This is the Social Security number listed for the Principal Investigator on the cover page (page 1) of the application (and again in the Key Personnel Table).

APPRCVD.

The APPRCVD field is the date listed on the cover page page 1) of the application denoting the date on which NHLBI received that application. This field enables us to distinguish between Round 1 and Round 2 submissions from the same institution.

REVSCORE.

This field contains the priority score assigned by the Grant Review Panel to a given application.

PMBSITE

The **PMBSITE** table contains a record for each performance site proposed in a POEMB application. This table is linked to **PMBAPP** (the table of POEMB applications) and **PMBPERS** (the table of individuals cited in POEMB applications and progress reports) through the fields ACODE and ICODE. Data fields in this table include ACODE, SITE#, ICODE, INST, DEPT1 and DEPT2, standard address information (ADD1, ADD2, CITY, STATE, ZIP), and TEL#.

ACODE.

In **PMBSITE**, the ACODE refers back to the **PMBAPP** table and denotes the application in which a given institution was listed as a "performance site."

SITE#.

SITE# is a Battelle-assigned number that denotes the order in which a given institution was listed as a site in a given application. SITE# alone is not a unique identifier field, although in combination with ACODE it forms a unique identifier, the ICODE (see below). Every application contains an institutional site assigned the number zero (0 = mailing address) and possibly additional performance sites (assigned the numbers 1, 2, 3 ...), as well. If one of the performance sites listed was identical to the mailing address, it was skipped (i.e., not assigned a number).

ICODE.

ICODE is a composite Battelle-assigned number formed from the ACODE (application code) and the SITE# separated by a decimal point. The ICODE is a unique identifier, referring to a specific institution, department, and address *as listed* either under "mailing address" or "performance site" on the cover page (page 1) of a given application.

It is possible for a given institution to be listed more than once in **PMBSITE** — if the same institution applied in both Round 1 and Round 2, for example, and listed some of the same performance sites. However, even if an institution was listed more than once in **PMBSITE**, the ICODE for each record is always unique and permits linkage back to the application in which the site was listed.

INST.

The field INST contains the institutional site as listed either in the mailing address or in the enumeration of performance sites on the cover page (page 1) of an application. When combined with the other address-related fields it produces a mailing address. This field is not related to the INST field in **PMBED**, which denotes the institution from which a person received a given academic degree.

DEPT1 and DEPT2.

The two DEPT fields (DEPT1 AND DEPT2) contain information on the professional school and/or department in which a given performance site is located and can be combined with other fields for a mailing address.

ADD1, ADD2, CITY, STATE, ZIP.

These are standard address fields which combine with INST and DEPT1 and DEPT2 to form a mailing address for a given institutional site.

TEL#.

The TEL# field is that listed for the PI on the cover page (page 1) of the application and is most closely associated with the mailing address. Any other (non-mailing address) performance sites in **PMBSITE** do not contain information in this field.

PMBPERS

PMBPERS is the core of the POEMB tables. It is a table of staff cited in the POEMB applications and progress reports.

PMBPERS can be linked to the application(s) in which an individual's name appeared (through ACODE to **PMBAPP**). The table can also be linked to **PMBSITE** (through the ICODE, indicating the institution with which an individual is associated). Both **PMBAPP** and **PMBSITE** are related to many records in **PMBPERS** (because each application and site have multiple staff associated with them), whereas each record in **PMBPERS** is related to only one record in **PMBAPP** or **PMBSITE** (because each record in **PMBPERS** is associated with only one application and one site). Battelle maintained this connection between the individual and the application in which s/he was cited until all data entry had been thoroughly checked and rechecked. However, prior to matching the POEMB database against the larger data sets, we eliminated duplicate person-records in **PMBPERS**, such as those that resulted from personnel duplications between Round 1 and Round 2 applications.

PMBPERS is also linked to **PMBED** (the educational background table). In this relationship, a

single **PMBPERS** record is linked to multiple records in **PMBED**, with each **PMBED** record containing information on one of the academic degrees held by an individual with a record in **PMBPERS**.

The fields in **PMBPERS** are ACODE, PERSON#, PCODE, three name fields (LAST, FIRST, MIDDLE), ICODE, SCODE, three initial fields (INIT1, INIT2, INIT3), SUFFIX, SS#, DOB, two title fields (TITLE1 and TITLE2), and three degree fields (DEG1, DEG2, and DEG3). The most critical fields for matching with the larger data sets are LAST, FIRST, MIDDLE, INIT1, INIT2, INIT3, SUFFIX, SS#, DOB, DEG1, DEG2, and DEG3.

ACODE.

ACODE is the Battelle-assigned application code number that links to **PMBAPP** and **PMBSITE**. Every individual in **PMBPERS** is linked through the ACODE to the application(s) in which her/his name appeared. This means that those cited in more than one application initially had more than one record in **PMBPERS**, although we subsequently identified and resolved these duplications before matching with the larger data sets was attempted.

PERSON#.

PERSON# is a Battelle-assigned number that denotes the order in which a given individual was listed in a given application. PERSON# is not a unique identifier field, although in combination with the ACODE it forms a unique identifier, the PCODE. For every application, an individual is assigned a PERSON# in sequential order starting with one (1 = Principal Investigator) and continuing with 2,3,4 ... for each additional key personnel. Potential trainees and support staff listed with the research component descriptions in the applications and trainees listed in progress reports are assigned the highest numbers in the series.

It is not, however, possible to determine where the breaks between key personnel, trainees, potential trainees, and support staff occur because the break-point numbers are different for each application. For this reason, we have added the SCODE (status code) field (see below), which clearly denotes an individual's role or status.

PCODE.

Like the ICODE (or institution code), the PCODE (or person code) is a unique Battelle-assigned identifier formed as a composite of ACODE (application code) and another field (PERSON#). This code identifies for each individual record in **PMBPERS** (1) the application in which that person was proposed and (2) the relative position in which that person was listed in the application.

LAST, FIRST, MIDDLE.

Initially, every person in **PMBPERS** was linked to the application in which her/his name appeared, which means that the same person (if bid on two different applications) could be entered under two different records. We subsequently identified and resolved such duplications before matching with the larger data sets was attempted.

ICODE.

If possible, the records in **PMBPERS** were associatd with a given performance site with which the person was associated in the application. Some applications, however, did not clearly define the association between key and support personnel and the various performance sites. In this case, the zero ICODE (mailing address) was used. In other cases, a given individual was associated with multiple performance sites (in which case again the zero ICODE was used).

SCODE.

The SCODE (status code) field denotes the level of responsibility proposed for an individual in a given application. These are:

Status 1	Principal Investigator (PI)
Status 2	Individual listed in the Key Personnel Table of the application, generally a PI of a proposed research component or otherwise an important investigator on the grant (Key Investigator).
Status 3	Individual listed in a proposed research component of an application or in a biosketch with a degree beyond the master's level but not appearing in a subsequent progress report (Potential Trainee).
Status 4	Individual listed in a proposed research component of an application or in a biosketch with a degree beyond the master's level and appearing in a subsequent progress report (Trainee).
Status 5	Individual cited in an application or progress report with a degree of master's level or below (Support Staff) or for whom no advanced degree is listed.

In some applications *all individuals proposed* appeared in the Key Personnel Table, even those without advanced degrees. Those with no advanced degrees were assigned Status 5 (Support Staff), despite the fact that they were listed in the Key Personnel Table. All others were assigned Status 2 (Key Personnel). Although it is conceivable that some recent MDs and PhDs might have been among them, we had no way of distinguishing between such Potential Trainees and "true" Key Personnel and so were obliged to accept the institution's designation of them all as Key Personnel.

Lack of evidence of an advanced degree was equated with lack of that degree, which means that some Status 3 or 4 individuals may have been misclassified as Status 5. Additional earned degree information received, either in a progress report or through institutional verification, occasionally allowed us to reclassify misclassified individuals. However, because of the possibility of misclassification, we included Status 5 individuals in all of our matching procedures.

Analyses that involved Investigators and their activities were run on a data set of Status 1 and Status 2 individuals (Principal Investigators and Key Personnel, respectively). Analyses that involved Trainees and their activities were run on a data set of Status 4 individuals.

A data set of Status 3 individuals could conceivably be used as a control group for Trainees, representing a set of recent MDs and PhDs not affected by work on the POEMB grant. Because they are either (1) recent advanced degree recipients proposed in non-successful applications or (2) recent advanced degree recipients proposed in a successful application but not appearing in a subsequent progress report, this group represents a set of Trainee-level individuals not materially affected by actual work on a POEMB grant.

INIT1, INIT2, INIT3.

The initial fields (INIT1, INIT2, INIT3) were filled in from information from the FIRST and MIDDLE name fields. We also scanned publication listings (such as those in the biosketches) to identify other possible ways that an individual might be cited in bibliographical and other databases. We noted any discrepancies indicating other possible forms of citation for a given individual in the DISCREPANCIES field of **PMBPERS** and also in the **PMBALT** file (see below).

SUFFIX.

This field contains name suffixes, such as Jr (Junior), Sr (Senior), or III (the Third).

SS#.

Whereas the SS# field in **PMBAPP** is the Social Security number of the Principal Investigator only, the SS# field in **PMBPERS** contains Social Security numbers for all staff. Key personnel generally have data in this field, since that information was usually available from the Key Personnel Table. Trainees (Status 4) may also have data in this field, available from progress reports. Potential Trainees (Status 3) have no data in this field because, in the applications, Social Security numbers were provided only for key personnel. If the name of a Potential Trainee from an application (Status 3) later appeared in a progress report by a successful POEMB applicant, complete with biosketch, that individual's status was changed to Status 4 (Trainee) and any new information available from the biosketch was added to the record.

DOB.

The DOB (date of birth) field was available for anyone for whom a biosketch was submitted as part of an application or progress report. That is, PIs, Key Personnel, and often Trainees. We do not have information on date of birth for any Potential Trainees or Support Staff, or for any Key Personnel or Trainees for whom no biosketch was submitted. In the Comments column of our institutional verification forms, we noted individuals for whom no biosketch had been submitted. A number of institutions sent us missing biosketches when they returned the verification forms. Data form these supplemental biosketches were sentered into **PMB INALT**.

TITLE1 and TITLE2.

The title fields (TITLE1 and TITLE2) contain an individual's academic title and/or position or role

in the proposed grant. Information for the title fields was culled from the Key Personnel Table (for PIs and key personnel), from the Trainee Table (for trainees on winning grants), and from the research component project descriptions in applications (for support staff).

DEG1, DEG2, and DEG3.

Degree fields (DEG1, DEG2, DEG3) contain information on the advanced academic degrees an individual has obtained (e.g., MD, PhD). Information for the degree fields was culled from the Key Personnel Table (for PIs and key personnel), from the Trainee Table (for trainees), and from the research component project descriptions in applications (for support staff).

Data Source (series of fields).

We created a series of fields to help identify the data sources used to obtain the information in each record of **PMBPERS**. The fields are designated as 01, 02, 03, 04, 05, 06, 07, referring to the last two digits in the application number cited on an application or progress report. That is, the field 01 refers to "1 P01 HL #####-01," which is the initial application by an institution; the field 02 refers to "1 P01 HL #####-02," which is the first progress report. These fields were marked with an "X" if an individual's name and additional information appeared in a given data source.

Staff proposed in non-successful applications have an X in the field 01 only. Staff proposed in successful applications may have multiple X's in the data source fields, assuming their names appear in an application and one or more progress reports.

This series of fields serves two purposes. First, we were able to double check discrepant information more easily because we knew precisely where the information came from. Second, in the case of successful grantee institutions, we have been able to track each individual's association with the POEMB grant. For example, an investigator's name/biosketch may appear in the application and two progress reports, but nowhere subsequent to that. This individual's association with the POEMB grant is therefore less long-lived than that of an investigator whose name appears in seven progress reports.

DISCREPANCIES.

This memo field was used to record discrepant information obtained from multiple data sources. We specify what the discrepancy is and the data sources in which the discrepant information appeared. We then used the information in this field to resolve subsequent matching problems and to construct the **PMBALT** table (see below).

ANECDOTES.

This memo field was used to record anecdotal information regarding the effect the POEMB program produced on the life and career of an individual associated with it. A trainee later elevated to investigator status was noted here, for example, or an investigator who wrote a letter attesting to the program's impact on the course of his career (if the letter was included with the application). Most information of this type was found in the progress reports.

COMMENTS.

This memo field was used for miscellaneous comments that did not clearly fit in either the DISCREPANCIES or the ANECDOTES field just described.

PMBALT

Often an individual was cited in one or more applications or progress reports, and occasionally we encountered discrepancies among data sources in terms of DOB, SS#, and the spelling of names. We had no way to determine which version was correct and so noted all versions in the DISCREPANCY field of **PMBPERS**. We asked the applicant institutions to review and verify summary printout from **PMBPERS**, but not all responded. Nor were all of the institutions that responded able to resolve all of the discrepancies we noted.

We then designed a separate table to house the alternatives that resulted from unresolved discrepancies. The fields in this table are self-explanatory as alternatives to important fields in **PMBPERS**.

LAST2	Alternative last name
FIRST2	Alternative first name
MIDDLE2	Alternative middle name
DOB-2	Alternative date of birth
SS#-2	Alternative Social Security number
INIT1-2	Alternative first initial
INIT2-2	Alternative second initial
INIT3-2	Alternative third initial
ACODE2 or ACODE3	Alternative application numbers (for those individuals proposed in more than one application)

For those individuals with hypenated last names, an alternative last name was automatically provided, whether or not data existed to suggest that the alternative was ever used. That is, the name Aguilar-Bryan was automatically given the alternative Bryan as a second possibility to match.

For those individuals whose names were listed with an initial followed by a full middle name (E. Michael Banks), indicating that they were better known by their middle name, an alternative first name was automatically provided (in this case, Michael), whether or not data existed to suggest that the alternative was ever used.

All matching procedures were run on both the **PMBPERS** and the **PMBALT** files to ensure that all viable alternatives were considered.

PMBINALT

This table is indentical in structure to **PMBALT** and was designed to house additional or alternative data reported by applicant institutions during institutional verification procedures. A summary report from this file was used in manual checking of questionable matches, and the file was used electronically for a second run against "no hit" records.

PMBED

The **PMBED** table summarizes the educational background and experience (if available) of those individuals with records in **PMBPERS**. The information was culled from the biosketches and was hence available only for PIs, Key Personnel, and Trainees for whom biosketches were submitted. The information was not available for Potential Trainees or Support Staff. The data fields INST (institution granting degree), DEGREE (academic degree granted), YEAR (year conferred), and FIELD (field of study in which degree was conferred) are self-explanatory. PCODE is the field that links the information in **PMBED** back to a given individual in **PMBPERS**.

The field of study in which a particular degree is awarded may be of interest in determining whether POEMB caused a subsequent shift in area of interest for that individual.

Note that the INST field in **PMBED** relates only to this table (denoting the institutions that awarded an individual's academic degrees) and bears no relationship to the INST field in **PMBSITE** (denoting a POEMB applicant institution).

Table 2A.1
Summary Table of Data Fields in POEMB Person-Based Paradox Tables

Field Name	Field Type	No. of Characters	Source of Data	Fields Linked to in Other Tables	Comments
PMBSITE *(table of all performance sites listed in POEMB applications)*					
Acode	Numeric		Battelle assigned	Acode in PMBAPP Acode in PMBPERS	Application code assigned by Battelle to each POEMB application submitted. Application codes 1-3 were assigned to the winning applications.
Site#	Numeric		Battelle assigned		Number assigned to each performance site proposed in a POEMB application. The number 0 is assigned to the site noted as the "mailing address." Additional numbers are assigned as needed to other performance sites listed. If the first performance site listed is identical to the "mailing address," then it is skipped (i.e., not assigned a number).
Icode	Numeric		Composite of Acode and Site# above	Icode in PMBAPP Icode in PMBPERS	Unique site identifier code for each site proposed in a POEMB application, formed as a composite of Acode and Site# (separated by a decimal point). That is, the second performance site (not counting the mailing address) listed on Application 12 would have the Icode 12.2.
Inst	Alphanumeric	50	POEMB application (page 1)		Name of institution listed as a performance site in a POEMB application.
Dept1 Dept2	Alphanumeric Alphanumeric	30 30	POEMB application (page 1)		Two fields for department of institution listed as a performance site in a POEMB application.
Add1 Add2	Alphanumeric Alphanumeric	30 30	POEMB application (page 1)		Two address fields for institution listed as a performance site in a POEMB application (if available).
City State Zip	Alphanumeric Alphanumeric Alphanumeric	25 2 12	POEMB application (page 1)		City, state, and zip fields for address of institution listed as a performance site in a POEMB application (if available).

Table 2A.1 (continued)
Summary Table of Data Fields in POEMB Person-Based Paradox Tables

Field Name	Field Type	No. of Characters	Source of Data	Fields Linked to in Other Tables	Comments
Tel#	Alphanumeric	14	POEMB application (page 1)		Telephone number, generally available for Principal Investigator/Primary Site only.
PMBAPP *(table of POEMB applications received by NHLBI)*					
Acode	Numeric		Battelle assigned	Acode in PMBSITE Acode in PMBPERS	Application code assigned by Battelle to each POEMB application submitted. Application codes 1-3 were assigned to the winning applications.
Icode	Numeric		Battelle assigned	Icode in PMBSITE	code for the primary performance site listed in a given POEMB application.
App#	Alphanumeric	16	POEMB application (page 1)		Application number (1 P01 HL ...)
SS#	Alphanumeric	11	POEMB application (page 1)	SS# in PMBPERS (for PI only)	Social Security number of Principal Investigator proposed in POEMB application.
Apprcvd	Date		POEMB application (page 1)		Date application received by NHLBI.
Revscore	Numeric		POEMB application review (page 1)		Priority score application received from Grant Review Panel.

Table 2A.1 (continued)
Summary Table of Data Fields in POEMB Person-Based Paradox Tables

Field Name	Field Type	No. of Characters	Source of Data	Fields Linked to in Other Tables	Comments
PMBPERS *(table of staff proposed in POEMB application)*					
Acode	Numeric		Battelle assigned	Acode in PMBAPP Acode in PMBSITE	Application code assigned by Battelle to each POEMB application submitted. Application codes 1-3 were assigned to the winning applications.
Person#	Numeric		Battelle assigned		Person number assigned to each person proposed in a POEMB application. The number 1 is assigned to the Principal Investigator. Additional numbers are assigned as needed, first to the staff members listed in the Key Personnel table, then to additional staff listed in an application under each component. Additional numbers were assigned to individuals cited in the progress reports of successful POEMB grantee institutions.
Pcode	Numeric		Composite of Acode and Person# above	Pcode in PMBED	Unique identifier code for each person proposed in a POEMB application or cited as a trainee in a progress report, formed as a composite of Acode and Person# separated by a decimal point. That is, the second person listed on Application 12 would have the Pcode 12.2. This would be the person listed immediately below the PI on the Key Personnel table in the application.
Last First Middle Suffix	Alphanumeric Alphanumeric Alphanumeric Alphanumeric	25 25 25	**POEMB application:** PI (page 1) Key (Key Personnel Table) Trainees and support staff (Proposed research component descriptions) **POEMB progress reports:** Trainees (Trainee Table)		Last name, first name, middle name (if available) of person proposed in POEMB application or cited as a trainee in a progress report. Also Suffix, for example, Jr (Junior), Sr (Senior), or III (the Third).

Table 2A.1 (continued)
Summary Table of Data Fields in POEMB Person-Based Paradox Tables

Field Name	Field Type	No. of Characters	Source of Data	Fields Linked to in Other Tables	Comments
Icode	Numeric		Battelle assigned	Icode in PMBSITE Icode in PMBAPP	Icode of person proposed in POEMB application or cited as a trainee in a progress report. An attempt is made to assign a given staff member to one of the performance sites listed in the application. If it is unclear which performance site the staff member is employed by or if the person is employed by more than one site, then the individual is assigned the Icode of the mailing address (*.0).
Scode	Numeric		Battelle assigned		Status code of person proposed in POEMB application (1 = Principal Investigator, 2 = Key Personnel listed in the Key Personnel Table of an application). A third status (3 = Potential Trainee) was created for individuals who were *not* listed as key personnel, but had degrees beyond the master's level and never appeared in a POEMB progress report; a fourth status for those cited as trainees in a progress report by a successful POEMB grantee (4 = Trainee); and a fifth status (5 = Support Staff) for those listed in an application or a progress report for whom we found no evidence of a degree beyond the master's level. Note: Status 3 and Status 4 individuals are likely at the same professional level, the difference between them being that Status 4 individuals are listed in a progress report as having *actually worked* on a POEMB grant.
Init1 Init2 Init3	Alphanumeric Alphanumeric Alphanumeric	1 1 1	POEMB application (from sources listed under Last, First, Middle above; otherwise from publication listings on biosketches, if available)		First (Init1) and middle (Init2) and a third (Init3) initial of person proposed in POEMB application. Publication listings in biosketches and elsewhere were cross-checked to ensure that these are the initials as used by the researcher in publishing articles. Any discrepancies were noted in the Discrepancies field and in PMBALT.

Table 2A.1 (continued)
Summary Table of Data Fields in POEMB Person-Based Paradox Tables

Field Name	Field Type	No. of Characters	Source of Data	Fields Linked to in Other Tables	Comments
SS#	Alphanumeric	11	POEMB application (Key Personnel table)	SS# in PMBAPP	Social security number of person proposed in POEMB application. Available only for PIs and Key Personnel, not for Status 3 (Potential Trainees) or Status 5 (Support Staff). Available for Status 4 personnel (trainees) cited in the more recent progress reports (i.e., those containing a formal Trainee Table similar to the Key Personnel Table in the applications).
DOB	Date		POEMB application (biosketches)		Date of birth for person proposed in POEMB application or cited as a trainee in a progress report. Available for those individuals for whom biosketches were included (generally not for Status 5 personnel). Also available for trainees cited in the more recent progress reports (i.e., those containing a formal Trainee Table).
Title1 Title2	Alphanumeric Alphanumeric	30 30	POEMB application (Key Personnel table; individual components for support personnel)		Two fields for academic title (or role/position on grant) for person proposed in POEMB application or cited as a trainee in a progress report.
Deg1 Deg2 Deg3	Alphanumeric Alphanumeric Alphanumeric	5 5 5	POEMB application (Key Personnel table; individual components for support personnel)		Three fields for degrees held (beyond undergraduate degrees) by person proposed in POEMB application or cited as a trainee in a progress report.
Data Source 01 02 03 04 05 06 07	Alphanumeric	1	An "x" in one of these fields indicates that the individual's name and other supplemental information appeared in a given data source.		Seven fields representing possible data sources for information contained in each record of PMBPERS: 01 (= original application), 02 (= first progress report), 03 (= second progress report), etc. The two digits (0#) correspond to the NIH grant/application number assigned to each documentary data source, that is 1 PO1 HL #####-01 is the number assigned by NIH to the original application; 1 P01 HL #####-02 to the first progress report by a successful applicant, etc.

Table 2A.1 (continued)
Summary Table of Data Fields in POEMB Person-Based Paradox Tables

Field Name	Field Type	No. of Characters	Source of Data	Fields Linked to in Other Tables	Comments
Discrepancies	Memo	Unlimited	Memo noted during data entry or during proofing of data.		Memo field for use to note discrepancies between the data obtained on a given individual from the various data sources. Most of this information also contained in PMBALT.
Anecdotes	Memo	Unlimited	Memo noted during data entry or during proofing of data.		Memo field for use to note anecdotal accounts of the POEMB program's impact on a given individual. Generally obtained from information contained in the progress reports. Examples: an individual is elevated from trainee status to investigator status; an investigator writes a letter attesting to the impact of the program on his career.
Comments	Memo	Unlimited	Memo noted during data entry or during proofing of data.		Memo field for miscellaneous notes on a given individual or her/his record.
PMBALT *(table of POEMB applications received by NHLBI)*					
Last2 First2 Midlde2 Suffix2	Alphanumeric Alphanumeric Alphanumeric	25 25 25	**POEMB application:** PI (page 1) Key (Key Personnel Table) Trainees and support staff (component descriptions) **Progress reports:** Trainees (Trainee Table)		Discrepant information was occasionally found on a person cited in more than one data source. PMBALT provided a way to include these alternatives in matching procedures. These fields contain alternative last name, first name, or middle name (if available) of a person proposed in a POEMB application or cited as a trainee in a progress report. Also an alternative Suffix, for example, Jr (Junior), Sr (Senior), or III (the Third).
DOB-2	Date		POEMB application (biosketches)		Alternative date of birth for person proposed in POEMB application or cited as a trainee in a progress report. Available for those individuals for whom biosketches were included (generally not for Status 5 personnel). Also available for trainees cited in the more recent progress reports (i.e., those containing a formal Trainee Table).

Table 2A.1 (continued)
Summary Table of Data Fields in POEMB Person-Based Paradox Tables

Field Name	Field Type	No. of Characters	Source of Data	Fields Linked to in Other Tables	Comments
SS#-2	Alphanumeric	11	POEMB application (Key Personnel Table)	SS# in PMBAPP	Alternative Social Security number.
Init1-2 Init2-2 Init3-2	Alphanumeric Alphanumeric Alphanumeric	1 1 1	POEMB application (from sources listed under Last, First, Middle above; otherwise from publication listings on biosketches, if available)		Alternative first (Init1), middle (Init2), and third (Init3) initial of person proposed in POEMB application. Publication listings were cross-checked in biosketches and elsewhere to ensure that these are the initials as used by the researcher in publishing articles.
Acode2 Acode3	Alphanumeric Alphanumeric				Alternative application numbers (for those individuals proposed in more than one application).
PMBINALT *(table ofalternative information on individuals cited in POEMB applications received by NHLBI that resulted from a verification process with POEMB applicant institutions)*					
[All fields identical to those in PMBALT]					This file is structurally identical to PMBALT. It contains alternative identifying information for use in matching procedures that was obtained from institutional respondents asked to review and verify basic information in a printout of PMBPERS.

Table 2A.1 (continued)
Summary Table of Data Fields in POEMB Person-Based Paradox Tables

Field Name	Field Type	No. of Characters	Source of Data	Fields Linked to in Other Tables	Comments
PMBED *(table of educational background of staff proposed in POEMB application)*					
Pcode	Numeric		Battelle assigned	Pcode in PMBPERS	Unique identifier code for each person proposed in a POEMB application or cited as a trainee in a progress report, formed as a composite of Acode and Person# separated by a decimal point. This code will link multiple records in the PMBED table to a given record in the PMBPERS table.
Inst	Alphanumeric	50	POEMB application (biosketch)	[Not related to Institution field in PMBSITE.]	Institution from which degree was obtained by person proposed in POEMB application or cited as a trainee in a progress report.
Degree	Alphanumeric	5	POEMB application (biosketch)		Degree obtained by person proposed in POEMB application or cited as a trainee in a progress report.
Year	Alphanumeric	4	POEMB application (biosketch)		Year degree conferred upon person proposed in POEMB application or cited as a trainee in a progress report.

2A.1 Training and Fellow File

The Trainee and Fellow File (TFF) is a subset of the 790,149 records (334,226 persons) in the 1994 NIH TFF archive file (ACC1P1Z.TFF1994). The TFF is an ASCII file and is identical in format to the 1994 NIH TFF archive file. This format is documented in the July 15, 1994 documentation of the 1994 NIH TFF archive file.

2A.2 Doctorate Records File

The Doctorate Records File (DRF) is a subset of the 1,103,196 records in the 1920-1993 Doctorate Records File. The DRF is a SAS data set and is identical in format to the 1920-1993 Doctorate Records File. This format is documented in the January 1995, National Resource Council documentation of the 1920-1993 Doctorate Records File.

Table 2A.2 Description of Variables in the CGAF/CRISP Analysis File

VARIABLE	ORIGIN	DESCRIPTION
ICDCAT	Calculated	ICD code classification; HL, non-HL
SETNO	ALLIDS,CGAF	Set Number (current)
ACTIVITY	CGAF	Activity (program code)
ICD	CGAF	Institute/Center/Division code
GRANTNO	CGAF	ICD + serial number
APPTYPE	CGAF	Type of Application
TOTREQ	CGAF	Dollars - Requested
TOTDOL	CGAF	Dollars - Total Awarded
PSCORE	CGAF	Priority Score
NPSCORE	CGAF	Normalized Priority Score
DSF	CGAF	DSF code
AWARD	CGAF	CGAF Award Indicator
DRFID	CGAF	DRF ID Number
PSTART	CGAF	Start Date 1
ACOUNT	Calculated	Number of amendments
CCOUNT	Calculated	Number of continuations
POEMB	Calculated	POEMB grant application indicator; 1=POEMB, 0=other
PCODE	PMBPERS, ALLIDS	Person code
ACODE	PMBPERS, PMBAPP	Application code
ICODE	PMBPERS	Institution code
SCODE	PMBPERS	Status code
LAST	PMBPERS	Last name
ACODE2	PMBPERS	Application code 2
ACODE3	PMBPERS	Application code 3
FIRST	PMBPERS	First name
MIDDLE	PMBPERS	Middle name
SUFFIX	PMBPERS	Suffix
INIT1	PMBPERS	First initial
INIT2	PMBPERS	Middle initial
INIT3	PMBPERS	Last initial
SS	PMBPERS	Social Security Number
DOB	PMBPERS	Date of birth
TITLE1	PMBPERS	Title 1
TITLE2	PMBPERS	Title 2
DEG1	PMBPERS	Degree 1
DEG2	PMBPERS	Degree 2
DEG3	PMBPERS	Degree 3

VARIABLE	ORIGIN	DESCRIPTION
APP	PMBPERS	
1	PMBPERS	Named in Application; X=named
2	PMBPERS	Named in 1st Progress Report; X=named
3	PMBPERS	Named in 2nd Progress Report; X=named
4	PMBPERS	Named in 3rd Progress Report; X=named
5	PMBPERS	Named in 4th Progress Report; X=named
6	PMBPERS	Named in 5th Progress Report; X=named
7	PMBPERS	Named in 6th Progress Report; X=named
LAST2	PMBPERS	Alternate last name
FIRST2	PMBPERS	Alternate first name
MIDDLE2	PMBPERS	Alternate middle name
INIT1_2	PMBPERS	Alternate first initial
INIT2_2	PMBPERS	Alternate middle initial
INIT3_2	PMBPERS	Alternate last initial
SS_2	PMBPERS	Alternate Social Security Number
DOB_2	PMBPERS	Alternate date of birth
OTHER_AL	PMBPERS	
ZDATE	Calculated	Date of first POEMB application
PAWARD	Calculated	POEMB application award indicator; 1=POEMB awarded, 0=POEMB not awarded
QUARTER	Calculated	Quarter of award relative to first POEMB application
INCGAF	Calculated	Indicates if person and quarter are in CGAF; 1=in CGAF, 0=not in CGAF
CC	Calculated	Crisp Classification; HL, MB, HL/MB, other

Table 2A.3 Description of Variables in the AHA Analysis File

Variable	Origin	Description
PCODE	PMBPERS, AHA1, AHA2, AHA3	Person code
SCODE	PMBPERS, AHA1, AHA2, AHA3	Status code
CATEGORY	AHA1, AHA2, AHA3	Matching category
CRITERIA	AHA1, AHA2, AHA3	Matching criteria
LASTNAME	AHA1, AHA2, AHA3	Last name
FRSTNAME	AHA1, AHA2, AHA3	First name
SSN	AHA1, AHA2, AHA3	Social Security Number
IDNO	AHA1	AHA ID number
IDSEQ	AHA1	AHA ID sequence
PROGRM	AHA1,AHA2, AHA3	Program type
ORIGIN	AHA1, AHA2, AHA3	The AHA division to which the application was submitted
STATUS	AHA1, AHA2, AHA3	Applicant status
INST	AHA1, AHA2, AHA3	Applicant Institution
FI	AHA1, AHA2, AHA3	First initial
LI	AHA1, AHA2, AHA3	Last initial
APPLID	AHA2,AHA3	Unique 8-digit number assigned to an application when it is submitted to AHA
MNAME	AHA2,AHA3	Middle name or initial
STATE	AHA2,AHA3	State of applicant Institution
YEAR	Calculated	Year of AHA award
AHA	Calculated	Indicates which AHA file the record originated in (1,2,3).
ACODE	PMBPERS, PMBAPP	Application code
ICODE	PMBPERS	Institution code
LAST	PMBPERS	Last name
ACODE2	PMBPERS	Application code 2
ACODE3	PMBPERS	Application code 3
FIRST	PMBPERS	First name
MIDDLE	PMBPERS	Middle name
SUFFIX	PMBPERS	Suffix
INIT1	PMBPERS	First initial
INIT2	PMBPERS	Middle initial
INIT3	PMBPERS	Last initial
SS	PMBPERS	Social Security Number
DOB	PMBPERS	Date of birth
TITLE1	PMBPERS	Title 1
TITLE2	PMBPERS	Title 2
DEG1	PMBPERS	Degree 1
DEG2	PMBPERS	Degree 2
DEG3	PMBPERS	Degree 3

Variable	Origin	Description
APP	PMBPERS	
_1	PMBPERS	Named in Application; X=named
_2	PMBPERS	Named in 1st Progress Report; X=named
_3	PMBPERS	Named in 2nd Progress Report; X=named
_4	PMBPERS	Named in 3rd Progress Report; X=named
_5	PMBPERS	Named in 4th Progress Report; X=named
_6	PMBPERS	Named in 5th Progress Report; X=named
_7	PMBPERS	Named in 6th Progress Report; X=named
LAST2	PMBPERS	Alternate last name
FIRST2	PMBPERS	Alternate first name
MIDDLE2	PMBPERS	Alternate middle name
INIT1_2	PMBPERS	Alternate first initial
INIT2_2	PMBPERS	Alternate middle initial
INIT3_2	PMBPERS	Alternate last initial
SS_2	PMBPERS	Alternate Social Security Number
DOB_2	PMBPERS	Alternate date of birth
OTHER_AL	PMBPERS	
ZDATE	Calculated	Date of first POEMB application
PAWARD	Calculated	POEMB application award indicator; 1 = POEMB awarded, 0 = POEMB not awarded
B4AFTR	Calculated	Indicator of AHA award before (-1), during (0), or after (1) POEMB award year
AAWARD	Calculated	Indicates if record was an AHA award (1)

Table 2A.4 Description of Variables in the Medline Analysis File

Variable	Origin	Description
TA	Medline	Journal title
TITLE	Medline	Article title
PCODE	Medline, PMBPERS	Person code
JCAT	Calculated	Journal classification
HI	Calculated	Mesh code indicates heart
HLI	Calculated	Mesh code indicates heart/lung
HMI	Calculated	Mesh code indicates heart/molecular biology
LI	Calculated	Mesh code indicates lung
MI	Calculated	Mesh code indicates molecular biology
OI	Calculated	Mesh code indicates other
YMPUB	Calculated	Month and year of publication
ACODE	PMBPERS, PMBAPP	Application code
ICODE	PMBPERS	Institution code
SCODE	PMBPERS	Status code
LAST	PMBPERS	Last name
ACODE2	PMBPERS	Application code 2
ACODE3	PMBPERS	Application code 3
FIRST	PMBPERS	First name
MIDDLE	PMBPERS	Middle name
SUFFIX	PMBPERS	Suffix
INIT1	PMBPERS	First initial
INIT2	PMBPERS	Middle initial
INIT3	PMBPERS	Last initial
SS	PMBPERS	Social Security Number
DOB	PMBPERS	Date of birth
TITLE1	PMBPERS	Title 1
TITLE2	PMBPERS	Title 2
DEG1	PMBPERS	Degree 1
DEG2	PMBPERS	Degree 2
DEG3	PMBPERS	Degree 3
APP	PMBPERS	
1	PMBPERS	Named in Application; X=named
2	PMBPERS	Named in 1st Progress Report; X=named
3	PMBPERS	Named in 2nd Progress Report; X=named
4	PMBPERS	Named in 3rd Progress Report; X=named
5	PMBPERS	Named in 4th Progress Report; X=named
6	PMBPERS	Named in 5th Progress Report; X=named
7	PMBPERS	Named in 6th Progress Report; X=named

Variable	Origin	Description
LAST2	PMBPERS	Alternate last name
FIRST2	PMBPERS	Alternate first name
MIDDLE2	PMBPERS	Alternate middle name
INIT1_2	PMBPERS	Alternate first initial
INIT2_2	PMBPERS	Alternate middle initial
INIT3_2	PMBPERS	Alternate last initial
SS_2	PMBPERS	Alternate Social Security Number
DOB_2	PMBPERS	Alternate date of birth
OTHER_AL	PMBPERS	
ZDATE	Calculated	Date of first POEMB application
PAWARD	Calculated	POEMB application award indicator; 1 = POEMB awarded, 0 = POEMB not awarded
QUARTER	Calculated	Quarter of publication relative to first POEMB application
INMED	Calculated	Indicates if record was in Medline (1)

Table 2A.5 ICD Codes Used to Classify CGAF Applications as Heart/Lung Related

ICD CODE	Description
HL	National Heart, Lung, and Blood Institute (NHLBI)
HV	Division of Heart and Vascular Diseases (NHLBI)
HB	Division of blood diseases & Resources (NHLBI)
HR	Division of Lung Diseases (NHLBI)
HI	Division of Intramural Research (NHLBI)
HO	Office of the Director (NHLBI)
HC	Division of Epidemiology and Clinical Applications (NHLBI)

Table 2A.6 CRISP Descriptors by Category

Category	CRISP Descriptors
Heart/Cardiac/ Cardiovascular	567-586 1390-1399
Lung	2591-2619 Excluding 2591 (6506) Cellular respiration 2017 (5738) Neoplasms of the respiratory system 0368 (2354) Bacterial Disease - Actinomycetales infection 0368 (2498) Tuberculosis 1525 (1998) Hypersensitivity, Respiratory Hypersensitivity 1525 (2157) Hypersensitivity, respiratory hypersensitivity - asthma 3130 (5811) Viruses, respiratory viruses
Molecular Biology	0944 (5919) Diseases - molecular level studies 1945 -1947 Excluding 1946 (1804) Molecular condensations - polymers 1255 (3896) Genetics, biochemical genetics - molecular genetics 1255 (4314) Genetics - molecular cloning 2001 (2389) Neoplasms characteristics - molecular level studies
Other	All remaining descriptors

Table 2A.7 Criteria for MeSH Code Categorization of Journal Articles.

Category	Criteria
Cardiovascular/ Heart (HI)	any MeSH terms containing the words heart, coronary or cardiovascular, or the prefix cardio- (except for cardiopulmonary, which was classified as H/L) were classified as "heart" (HI) any MeSH terms that referred to the following: the anatomy of the heart (atrium/atrial, ventricle/ventricular, pericardium/pericardial, etc.) or vascular system (arteries/arterial, arterioles, blood vessels, etc.) cardiovascular diseases (arterial occlusive diseases, artherosclerosis, ischemia, myocarditis, etc.) medical or surgical procedures related to the cardiovascular system (angiography, angioplasty, atherectomy, etc.) blood pressure (hypertension, hypotension, systole, diastolic, etc.) chemicals, drugs, and endogenous substances that affect the cardiovascular system (angiotensins, anti-arrhythmia agents, antihypertensive agents, heparin, etc.)*
Cardiovascular/ Pulmonary (HI/LI)	MeSH terms referring to both the cardiovascular system and the pulmonary system (pulmonary circulation, pulmonary artery, pulmonary hypertension, etc.)
Lung (LI)	any MeSH terms containing the words respiratory, pulmonary, or lung, or the prefix broncho- were classified as "lung" (with the exception of H/L-designated terms). any MeSH terms that referred to the following: chemicals, drugs, and endogenous substances that affect the respiratory/pulmonary system (surfactant, bronchodilator agents, etc.) medical or surgical procedures related to the respiratory/pulmonary system (bronchoscopy, bronchoaveolar lavage, exogenous surfactant replacement therapy, etc.) pulmonary diseases (asthma, alveolitis, pleural diseases, etc.) the anatomy of the pulmonary system (alveoli, bronchi, etc.)

Category	Criteria
Molecular Biology (MI)	MeSH terms referring to cell lines, inbred strains of laboratory animals and microbes were classified as "molecular biology" due to their implicit use in molecular/biologic investigations. any MeSH terms containing the acronyms DNA and/or RNA and followed by the category abbreviations molecular biology (mb), microbiology (mi), virology (vi), and/or genetics (ge).* any MeSH terms that referred to the following: molecular biologic techniques and the technology employed in these techniques (alternative splicing, autoradiography, affinity chromatography, mutational analysis, restriction enzymes, sequence analysis, transfection, etc.)
"Other"	any MeSH term relating to all other (excluding pulmonary and cardiovascular) major organs or organ systems was classified as "other" as were any terms that did not fit the above-mentioned criteria (brain mapping, retinoblastoma, zoonoses, etc.)

* The Medical Subject Headings Annotated Alphabetic List was used in this classification.

Appendix 3A

Indices of Data Items Used

Appendix 3A

DRF INDEX OF DATA ITEMS USED					
Nbr	Variable Name	Type	Length	Position	Label
6	IDNUMBER	Character	7	12	IDNUMBER
41	PHDFIELD	Character	3	115	PHD MAJOR

TFF INDEX OF DATA ITEMS USED						
Variable Name	Source	Start Position	Length	Type	Analytic File Variable Name	Field Description on Page[a]
Set Number (generated)	QRC Generated	1	8	Alpha	SETNO	1
PHS ID Number:	IMPAC x1, TAF x1	259 - 277	1	Alpha	APPTYPE	63
Activity/Program Code (edited)	IMPAC x2, TAF x2	268	2	Alpha	ACTIVITY	70-71
Tenure - Entry Date (edited)	IMPAC x20, TAF x180	282	6	Num	TSTART	74
Tenure - Termination Date (edited)	IMPAC x90, TAF x181	288	6	Num		75
Variable Name	Source	Tape Position	Variable Type		Analytic File Variable Name	Page
Award Indicator	QRC Generated	356	Num		Award	97

[a] See *Documentation: 1994 Trainee and Fellow File*

CGAF INDEX OF DATA ITEMS USED					
Variable Name	Source	Tape Position	Variable Type	Analytic File Variable Name	Field Description on Page[a]
Institute (NIH) Identification Number	Pre-1962 Source IMPAC Item x3	2-15	Alpha	GRANTNO (4-8) SUPPYR (10-11) GRNSUFX (12-15)	4-6
Award Indicator	QRC generated	510	Numeric	AWARD	128
Start Date 1	IMPAC Item x41	284-289	Numeric	PSTART	72
End Date 1	IMPAC Item x21	290-295	Numeric	PEND	73
Set Number (Current)	QRC generated	470-477	Character	SETNO	117
Activity (Program Code)	Pre-1962 Source IMPAC Item x2	16-18	Alpha	ACTIVITY	7

[a] See *Documentation: 1994 Consolidated Grant Applicant File*

Appendix 6A

Interview Guides

6A.1 Interview Guide
POEMB Principal Investigators

Section I. Background Information

1. Briefly describe your involvement with the Program of Excellence in Molecular Biology (POEMB) at your institution.

[Prompt interviewee to address the following aspects of their involvement]

- Positions and/or titles (also obtained from Renewal Applications/Progress Reports).
- Length of time of involvement (also obtained from Renewal Applications/Progress Reports).
- Degree of involvement, i.e., full-time, part-time, etc (Compare to "% effort on project" reported in the Renewal Applications/Progress Reports).
- Component or aspect of the POEMB participated in (Compare to "% effort on project" reported in the Renewal Applications/Progress Reports).

Section II. Value of POEMB--Advantages and Disadvantages

1. Briefly explain why you decided to organize a response to the initial POEMB RFA.

Probe:

a. What specific **aspects** of the POEMB concept prompted you to organize an application and why?

[Prompt interviewee to address the five key features of POEMB]

- *The extended seven year award period.**
- *Cross-disciplinary recruitment and collaboration.**
- *Opportunity to enhance facilities and resources.**
- *Broad degree of research freedom--ability to implement innovative approaches.*
- *Support for new investigators.*

* These have traditionally been **highly** desirable opportunities. It would be interesting to see the individual weights that the interviewees place on the five categories.

b. Which of these **aspects** would you consider the most important in prompting you to organize/submit a POEMB application and why?

2. Did you anticipate any **disadvantages** that might be associated with the POEMB approach?

a. If so, what were they and have they proved to be a problem?

Section III. Accomplishments

1. What is the most **significant accomplishment** (or accomplishments) that has come out of your grant?

   [Prompt interviewee with highlights from the respective progress reports and renewal applications. List these highlights in a separate companion document. Highlights of accomplishments will be unique to each program. The companion document(s) should be tailored for each interview.]

2. Did the POEMB concept **contribute** to this accomplishment (or accomplishments), and if so **how**?

   Prompt:

   a. What role did the key **features** of the POEMB concept play, i.e.,

   - *The extended seven year award period.*
   - *Cross-disciplinary recruitment and collaboration.*
   - *Opportunity to enhance facilities and resources.*
   - *Broad degree of research freedom--ability to implement innovative approaches.*
   - *Support for new investigators.*

   b. Can you list these features in order of relative importance to the accomplishments achieved in your program?

3. Could such an accomplishment have been achieved outside of POEMB?

   a. Why or why not?

   b. Are you aware of other programs comparable to POEMB? If so, what is your general assessment of them?

4. Did the POEMB **concept or structure** hinder the program's research in any way?

   If yes, then

   a. What aspects of the POEMB concept were a hindrance and why?

   b. What specific aspects of your research program were adversely affected?

   c. How was this dealt with ?

      e.g.; remedial action, change in research focus or direction, personnel changes, *etc.*

   d. Did the Program/research benefit as a result of the changes?

5. Do you believe that, thus far, your program has been successful in achieving **NHLBI's primary objectives**? Explain.

   [**Prompt** interviewee to address each of the two primary objectives.]
   - *To foster utilization of molecular biology approaches in important research areas (cardiovascular and pulmonary) within the mission of the NHLBI (technology/research development and application).*
   - *To provide opportunities for investigators who have the potential for independent research careers to become skilled in the experimental strategies and techniques of molecular biology and their application to research relevant to the mission of NHLBI (training, dissemination and fostering the NHLBI mission).*

   a. Can you give an example of each objective within your program?

   b. Was there an emphasis on one objective more than the other, or were the program efforts and resources equally distributed?

6. Are there any ways that **your** program did not succeed in achieving the primary objectives?

   If Yes, then
   a. Which of the objectives were more problematic?

   b. Why?

   c. What could have been done differently?

      e.g., administratively, logistically, personnel, research focus, *etc.*

7. To what extent has your program been successful in achieving its **own** specific objectives?

   a. Explain.

   [Prompt interviewee to address each of the specific objectives gleaned from the applications and progress reports. These should be provided in a companion document. Objectives and components will be unique to each program. The companion document(s) should be tailored for each interview.]

8. To what extent did your program not succeed in achieving its specific objectives?

   a. Which objectives were most problematic?

   b. Why?

   c. What could have been done differently?

e.g., administratively, logistically, personnel, research focus, *etc.*

Section IV. POEMB Network

1. Were the individual programs able to establish and develop an overarching POEMB **network** in which interaction, collaboration, and the exchange of information occurred?

If so, then:

a. Please describe the network. What constituted this network?

e.g., personal contacts, Internet communication, conferences, collaboration in publishing, co-presentations at meetings, technology/personnel exchanges, cross-training of personnel, *etc.*

b. In what ways was your program involved in creating and maintaining this network?

e.g., inviting speakers from other POEMB programs, providing facilities for collaboration or meetings, *etc.*

c. Was this network beneficial to your specific program, and why or why not?

e.g., increased technology access, other sources of qualified trainees or investigators, *etc.*

Section V. Recommendations

1. What **recommendations** would you make to NHLBI for the future of POEMB and the pursuit of its primary objectives?

POEMB Objectives:

- *To foster utilization of molecular biology approaches in important research areas (cardiovascular and pulmonary) within the mission of the NHLBI (technology/research development and application).*
- *To provide opportunities for investigators who have the potential for independent research careers to become skilled in the experimental strategies and techniques of molecular biology and their application to research relevant to the mission of NHLBI (training, dissemination and fostering the NHLBI mission).*

6A.2 Interview Guide
POEMB Associate Investigators

Section I. Background Information

1. Briefly describe your involvement with the POEMB.

   [Prompt interviewee to address the following aspects of their involvement]
   - Positions and/or titles (also obtained from Renewal Applications/Progress Reports).
   - Length of time of involvement (also obtained from Renewal Applications/Progress Reports).
   - Degree of involvement, i.e., full-time, part-time, etc (Compare to "% effort on project" reported in the Renewal Applications/Progress Reports).
   - Component or aspect of the POEMB participated in (Compare to "% effort on project" reported in the Renewal Applications/Progress Reports).

Section II. Value of POEMB--Advantages and Disadvantages

1. **Why** did you participate in the response to the initial POEMB RFA?

   Probe:
   a. What specific **aspects** of the POEMB concept prompted you to be involved in this response and why?

   [Prompt interviewee to address the five key features of POEMB]
   - *The extended seven year award period.**
   - *Cross-disciplinary recruitment and collaboration.**
   - *Opportunity to enhance facilities and resources.**
   - *Broad degree of research freedom--ability to implement innovative approaches.*
   - *Support for new investigators.*

   * These have traditionally been **highly** desirable opportunities. It would be interesting to see the individual weights that the interviewees place on the five categories.

   b. Which of these **aspects** would you consider the most important in prompting you to participate in the submission of a POEMB application and why?

2. What did you see as the **disadvantages** of the POEMB approach?

   a. Has your opinion changed?

Section III. Accomplishments

1. What is the most **significant accomplishment** (or accomplishments) that has come out of the program, and how is your component related to it?

   [Prompt interviewee with highlights from the respective progress reports and renewal applications. List these highlights in a separate companion document. Accomplishments will be unique to each program and component. The companion document(s) should be tailored for each interview.]

2. How did the POEMB approach **contribute** to this accomplishment?

   Prompt:
   a. What role did the key **features** of the POEMB concept play, i.e.,
   - *The extended seven year award period.*
   - *Cross-disciplinary recruitment and collaboration.*
   - *Opportunity to enhance facilities and resources.*
   - *Broad degree of research freedom--ability to implement innovative approaches.*
   - *Support for new investigators.*

   b. Were some of the key features more important than others in contributing to the Program's accomplishments? Why?

3. Would it have been possible for these accomplishments to be achieved outside of POEMB?

   a. Why or why not?

   b. Are you aware of other programs similar to POEMB? If so, how do they compare?

4. Did the POEMB **concept or approach** hinder your component's research in any way?

   If yes, then
   a. In what way was the POEMB concept or approach a hindrance and why?

   b. What specific aspects of your research component were adversely affected?

   c. How did you and your colleagues deal with it?

      e.g.; remedial action, change in research focus or direction, personnel changes, *etc.*

   d. Did your component benefit as a result of the changes?

5. Has the program has been successful in achieving **NHLBI's primary objectives**? Explain.

5.0 Journal Article Publication Activity

5.0 Journal Article Publication Activity

The publishing of journal articles is another dimension of research activity that was examined in the evaluation of the POEMB program. Journal articles provide evidence related to the overall research activity of individual investigators as well as the substantive area and focus of those activities. By examining the publication patterns of individuals associated with the successful and unsuccessful POEMB applications, both before and after the award of the three POEMB grants, we were able to assess the possible effect of the POEMB on those individuals' research activities. Specifically, we sought to determine whether the POEMB encouraged researchers in the successful applicant group to apply molecular biologic techniques to cardiovascular and pulmonary disease research problems. Also conceivable was a similar effect among unsuccessful applicants, which might suggest that the thought and collaborative effort involved in preparing a POEMB grant may have carried over into research activity.

In this chapter, we focus on comparisons between the journal article publication activity of the researchers in the two POEMB applicant groups. In addition, we describe the publication activity of those researchers who left the POEMB before the end of the program in an effort to determine whether these individuals continued to pursue research that involved the application of molecular biology to fields of cardiovascular and pulmonary research. The data and methodology used in this assessment are described first, followed by a discussion of the findings.

This chapter is subdivided into sections that discuss analyses focused on: (1) total publication activity, (2) publication activity as measured by journal title descriptor categories, (3) "switching" in terms of journal title descriptor categories, (4) publication activity as measured by MeSH term categories, (5) "switching" in terms of MeSH term categories, and (6) "switching" in terms of journal title descriptor categories combined with MeSH term categories.

5.1 Methods

The **Medline** baseline file was used for all analyses of journal publication activity reported here. This file contains journal citations and abstracts from medical and scientific journal articles obtained from a **Medline** search carried out by the National Library of Medicine (NLM). Using a file developed by Battelle, NLM searched **Medline** for articles published by people in the **PMBPERS** file based upon

variations of the author's first and last name. Journal articles were assigned to quarters relative to each person's first POEMB application. Quarter "0" was defined such that a person's first POEMB application occurs at the midpoint of the quarter. For the purposes of this analysis, quarter "0" was included in the period before POEMB application ("pre-" POEMB application) and only the period from quarter -16 to quarter +16 (i.e., four years before and four years after POEMB application date) was considered for analysis. A detailed description of this data set can be found in Chapter 2 of this report.

Using the 1996 version of *Ulrich's International Periodicals Directory*, publications by researchers in both POEMB applicant groups were classified into eight categories based upon journal title. These title categories included: cardiovascular/heart (H), respiratory/lung (L), molecular/microbiology (M), heart and/or lung (HL), heart and molecular biology (HM), lung and molecular biology (LM), "cross-cutting" (X), and "other" (O).

The complete list of journal titles from the **Medline** baseline file was checked against this reference text, and key *Ulrich's* subject headings were used to classify the titles into the above-mentioned categories. This journal title classification scheme is summarized in Table 5.1. It should be noted that the Battelle journal title categories heart and molecular biology (HM) and lung and molecular biology (LM) each contain only one journal title. The category "cross-cutting" was created to classify those journals that span all of the previous categories. An exhaustive list of journals classified as cross-cutting is provided in Table 5.1.

Searches were conducted on the World Wide Web for information on any journals not found in *Ulrich's International Periodicals Directory*. Information from the aims, scope, and/or audience statements for all journals not listed in this reference text was available on Web Pages maintained by the publishers of these journals. This information was used to classify these journals into the categories listed above.

Table 5.1 Journal Title Classification Scheme Used in Categorizing the Medline Baseline File Journals

Ulrich's International Periodicals Directory Subject Heading	Battelle's Journal Title Category	Examples of Journal Titles*
Cardiovascular Diseases	cardiovascular/heart (H)	• *Journal of the American College of Cardiology* • *American Heart Journal* • *American Journal of Cardiology*
Respiratory Diseases	respiratory/lung (L)	• *Experimental Lung Research* • *Lung* • *American Review of Respiratory Disease*
Cardiovascular Diseases and Respiratory Diseases	heart and/or lung (HL)	• *Annals of Thoracic Surgery* • *Journal of Thoracic Imaging* • *Journal of Thoracic and Cardiovascular Surgery* • *Chest* • *Clinics in Chest Medicine*
Genetics	molecular/microbiology (M)	• *American Journal of Medical Genetics* • *American Journal of Human Genetics* • *Gene*
Microbiology	molecular/microbiology (M)	• *Journal of Virology* • *Journal of Bacteriology* • *Molecular and Cellular Biology*
Cardiovascular Diseases and Genetics	heart and molecular/ microbiology (HM)	• *Journal of Molecular and Cellular Cardiology*
Respiratory Diseases and Genetics	lung and molecular/ microbiology (LM)	• *American Journal of Respiratory Cellular and Molecular Biology*
All Previous Subject Headings	"cross-cutting" (X)	• *American Journal of Medicine* • *American Journal of Medical Science* • *Annals of Internal Medicine* • *Journal of Clinical Investigation* • *Lancet* • *New England Journal of Medicine* • *Nature* • *Proceedings of the National Academy of Science (USA)* • *Science*
None of the Previous Subject Headings	"other" (O)	• *Scanning Microscopy* • *Placenta* • *Journal of Bone Mineral Research*

* Titles listed under the categories H, L, M, and O are merely examples. However, the lists of titles included under the categories HL, HM, LM, and X are exhaustive.

Journal articles in the **Medline** baseline file were also characterized by their associated medical subject headings (MeSH terms). These categories include: heart (HI), lung (LI), molecular biology (MI), heart/lung (HLI), and "other" (OI). See Appendix 2A for a detailed description of the criteria used for classifying MeSH terms. Because each journal article is associated with one or more subject headings, there are 14 possible MeSH term categories that can be associated with an individual article. These categories include:

HI only	LI only	MI only
HI/MI	LI/MI	HI/LI/MI
"other" only	HI/ "other"	LI/ "other"
MI/ "other"	HI/LI/ "other"	HI/MI/ "other"
LI/MI/ "other"	HI/LI/MI/ "other"	

It should be noted that all articles in the **Medline** file were associated with at least one "other"-designated MeSH term.

As a third way of characterizing journal articles, we combined the journal title category descriptors with the MeSH term category descriptors. This allowed for the examination of the **Medline** data in terms of the percentage of investigators who published articles in all journal title categories along with MeSH term-designated categories of interest (HI/LI, MI, and HI/LI/MI).

In an approach consistent with the goal of assessing whether the POEMB affected the research direction of individuals in the successful POEMB applicant group, these analyses focus mainly on the investigator as the unit of analysis. The *dependent variables* considered in these analyses obtained directly from the **CGAF/CRISP** baseline file include (1) total publications, (2) the distribution of journal title categories, and (3) the distribution of MeSH term categories. Those dependent variables that were constructed using data from the **Medline** baseline file include (4) the ratios of journal title descriptor categories to total publications; (5) the ratio of MeSH term categories to total publications; (6) having published one or more journal articles versus having published no articles; (7) having published one or more articles in H, L, M, H/M, L/M, or "other"-designated journals versus no publications; (8) having published one or more articles with HI/LI, MI, and/or HI/LI/MI MeSH term designations; and (9) the percentage of investigators that published articles in HI/LI journals with MI MeSH terms, articles in MI

journals with HI/LI MeSH terms, articles in X journals with HI/LI/MI MeSH terms and articles in O journals with HI/LI/MB MeSH terms.

The analyses discussed in this chapter focus on the following measures of the differences in the publication activity of researchers in the two POEMB applicant groups:

- Aggregate number of total journal articles published in the pre-POEMB award, post-POEMB award, and total periods.
- Mean number of total articles and mean number of articles published in H-, L-, M-, X-, and O-designated journals in the pre-POEMB award, post-POEMB award, and total periods.
- Mean proportion of total publications appearing in H-, L-, M-, X- and O-designated journals in the pre-POEMB award, post-POEMB award, and total periods.
- Mean number of articles associated with HI/LI, MB, and HI/LI/MB MeSH terms in the pre-award, post-award, and total periods.
- Mean proportion of articles associated with HI/LI, MB, and HI/LI/MB MeSH terms in the pre-award, post-award, and total periods.
- Visual presentation of each dependent variable over time (17 quarters before and 16 quarters after POEMB application), with trend lines to fit the data.
- Repeated measures analysis of the above-noted dependent variables.
- Transition probabilities from all possible pre-POEMB states (journal title categories, MeSH term categories, and co-occurring MeSH term and journal title categories) to all possible post-POEMB states.
- Proportional differences in terms of "switches" (i.e., changes from certain journal title categories in the pre-POEMB period to other key categories in the post-POEMB period; changes from certain MeSH term categories in the pre-POEMB period to other key categories in the post-POEMB period; changes from certain journal title/MeSH term categories in the pre-POEMB period to other key categories in the post-POEMB period).

All differences were examined by pooling all status types (SCODEs 1, 2, 3, and 4), and then comparisons were made between the senior investigators (SCODEs 1 and 2) in the two POEMB applicant groups as well as between trainees and potential trainees (SCODEs 3 and 4 for successful applicants are compared to 3 for unsuccessful applicants). As with applications and awards, it was important to break out senior investigators and trainee-level researchers because the program was expected to affect senior investigators and trainees differently on criterion outcomes, and also because the percentage of senior investigators was

far lower in the successful than in the unsuccessful applicant group. The highly unbalanced distributions of senior investigators versus trainees across the two applicant groups (13 percent versus 63 percent) can lead to misleading results when the senior investigators and trainee-level researchers are pooled. In the most extreme case — a phenomenon known as Simpson's paradox (Bickel et al., 1975) — the data can show a consistent result across both subgroups (e.g., both senior investigators and trainees obtaining more publications on average in the successful applicant group), while the pooled data show the opposite result (all investigators obtaining fewer publications on average in the successful applicant group).

In all analyses discussed in this chapter, status 3s (potential trainees) and 4s (trainees cited in POEMB progress reports) in the successful POEMB applicant group are combined as "trainees". We recognize that conceptually there are no individuals in the unsuccessful POEMB applicant group that are truly comparable to the status 4s in the successful applicant group. In order to investigate the potential bias introduced by including the status 4s in the "trainee" comparisons, we also compared the equivalency of the status 3s and 4s in the successful POEMB applicant group on all outcomes of interest. Cases in which the exclusion of the status 4s may have made a difference in the results and interpretations are cited in the text of this chapter.

As in Chapter 4, visual comparisons over the entire pre-post period (longitudinal) are an important focus of this analysis. These comparisons are provided in the form of plots over time and trend lines to fit the plots.

In all cases where the data support it, statistical modeling was carried out on the full sample (senior investigators and trainees pooled) in the form of repeated measures analysis. All such analyses were conducted using PROC MIXED in SAS®, with POEMB modeled as a fixed effect, investigator as a random effect nested within POEMB, the dependent variables (separately) as polynomial functions of time (17 quarters before to 16 quarters after POEMB application), and the covariance structure of repeated observations determined empirically (compound symmetry or autoregressive). The repeated measures analysis permits us to statistically test the hypothesis that receiving a POEMB award influenced the dependent variables in question. It is superior to a simple comparison of pre-post gain scores because: (1) it uses many more observations (quarterly data), thereby enabling the detection of trends that would not be apparent from pre-post gain scores in which those observations are aggregated (e.g., quadratic effects of time), and (2) it is more statistically powerful. The particular repeated measures model used, a mixed model in which the covariance structure of the quarterly observations was determined empirically,

provides the most accurate probability levels for significance tests of model parameters (Littell et al., 1996).

The main (null) hypothesis tested in this section is that there are no (statistically significant) group by time interactions. Given that the POEMB's extended award period of seven years was adopted to enable POEMB researchers "to pursue new developments in the rapidly advancing field of molecular biology and to embark on the application of molecular biology to complex experimental systems and new experimental models that would require additional time to develop and use," we would expect that receiving this award would allow POEMB investigators to engage in longer range, more innovative research. This effect may manifest itself as a decrease in the total publication activity of the POEMB investigators during the post-POEMB application period — this might be because their research is more innovative and thus less likely to be published. Also, longer range research may take longer to yield results, which would slow publication. We would also expect that, for the successful applicants, receiving the POEMB may have led to one of two possible pre- to post-POEMB shifts in research focus as measured by journal title and MeSH term categories — a shift from a focus that is primarily on cardiovascular or pulmonary disease research to one that incorporates molecular biologic techniques in the study of these diseases, or a shift from a focus that is primarily on the application of molecular biologic techniques to nonheart/lung related studies to one that applies those techniques to fundamental cardiovascular or pulmonary research problems.

5.2 Findings

The findings section is divided into four subsections: (1) total publication activity, (2) journal title descriptors, (3) article descriptors or MeSH terms, and (4) journal title descriptor-MeSH term combinations.

5.2.1 Total Publication Activity

Table 5.2 presents the raw frequencies of journal articles published by both POEMB applicant groups by status (SCODE) during the period 17 quarters before and 16 quarters after POEMB application. Researchers in the successful POEMB applicant group (N=264) published a total of 1,586 journal articles

Table 5.2 Journal Articles Published by Researchers in the Successful and Unsuccessful POEMB Applicant Groups for the Period 17 Quarters Before and 16 Quarters After POEMB Application

Investigators from Successful POEMB Applicant Group	Journal Articles Published 17 Quarters Pre-POEMB Row %, Column %	Journal Articles Published 16 Quarters Post-POEMB Row %, Column %	Total
Status 1 and 2 (Status 1, N=4; Status 2, N=30)	281 32% 65%	611 68% 53%	892
Status 3 (N=32)	46 38% 11%	76 62% 7%	122
Status 4 (N=198)	106 19% 24%	466 81% 40%	572
Investigators from Unsuccessful POEMB Applicant Group			**Total**
Status 1 and 2 (Status 1, N=18; Status 2, N=365)	2079 30% 94%	4738 70% 91%	6817
Status 3 (N=220)	138 22% 6%	483 78% 9%	621
Total	2217	5221	

during this time period. A total of 7,438 journal articles were published by the individuals in the unsuccessful applicant group (N=603) during this same period. A total of 32 authors[1] contributed to the 892 journal articles published by senior investigators during the total period. The 122 articles published by the POEMB status 3s during the total period represent the work of 22 authors in this stratum. Ninety-eight of the 198 POEMB status 4s contributed to the 572 articles published by this stratum during the total period. A total of 336 senior investigators from the unsuccessful POEMB applicant group authored the 6,817 articles published during this same period. The 621 articles published by the potential trainees in the unsuccessful applicant group represents the work of a total of 136 authors from this stratum.

[1] "Author" here means any person listed as an author on the journal article. No distinction is made between first and other authors.

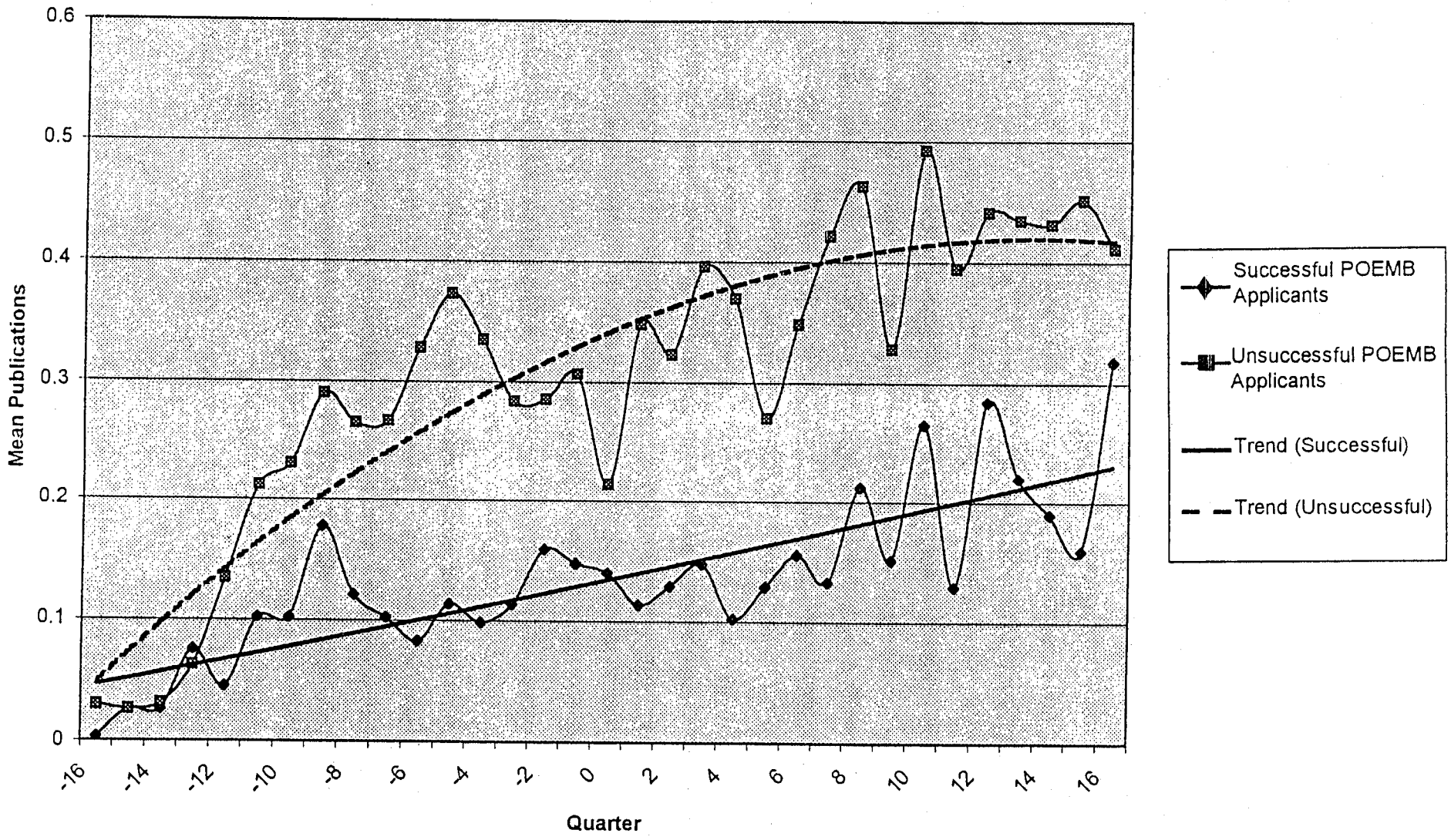
An examination of the data presented in Table 5.2 shows that all investigators, regardless of status or applicant group, increase in publication activity from the pre- to the post-POEMB period. The most apparent pre-post-POEMB patterns in publication activity are the almost four-fold increases in total number of articles published by the status 4s in the successful POEMB applicant group (106 to 466) and the status 3s in the unsuccessful group (138 to 483). Because similar increases are observed for these strata in both of the POEMB applicant groups, these changes in activity are most likely due to an effect of "maturation." "Maturation" in this sense refers to the fact that investigators will become more active in terms of publishing as their research career develops.

On average, unsuccessful POEMB applicants published more than twice as many journal articles tracked by **Medline** as did successful POEMB applicants. These relationships held over the total period, pre-POEMB period, and post-POEMB period, and each contrast was statistically significant ($p < 0.01$).[1] In addition, the analysis of gain scores showed these differences increased over the pre-post period ($p < 0.01$). Figure 5.1 plots these activities by quarter from 17 quarters prior to the POEMB application date to 16 quarters post, along with the least squares quadratic trend line through each time series. As shown, the trends reflect both the differences between groups and their relative increase over time. Examination of these same data by status showed that for the total period, POEMB senior investigators (SCODEs 1 and 2) had a higher mean number of publications per investigator than did the senior investigators in the unsuccessful POEMB applicant group. This difference was not statistically significant, however, due to low power (there were only 34 POEMB senior investigators). Both groups of senior investigators showed about the same amount of increase from pre- to post- with the successful POEMB applicants starting higher and ending higher. A comparison of the status 3s and 4s in the successful POEMB applicant group to the status 3s in the unsuccessful group showed that these two groups were similar in terms of their mean number of publications for all periods. The pattern observed here — unsuccessful applicants publishing higher mean numbers of articles overall, with successful applicants posting higher mean numbers of publications in each of the two subgroups — results from a phenomenon known as Simpson's paradox (Bickel et al., 1975). As described in the methods section, it stems from the highly unbalanced distributions of senior investigators versus trainees in the two POEMB applicant groups (i.e., 13 percent of senior investigators among successful applicants compared to 63 percent among unsuccessful applicants).

---

[1] Overall group differences were tested with t-tests. Where group variances were unequal ($p < 0.05$), the Cochran and Cox (1950) approximation was used.

Figure 5.1 Mean Number of Journal Publications Per Investigator, By Quarter



Examination of the gain score data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

The same data were examined using a mixed-model repeated measures analysis, with POEMB status as a fixed effect, investigator as a random effect nested within POEMB status, the dependent variables specified as polynomial functions of time, and the covariance structure of the repeated observations determined empirically. The fixed main-effect terms included in the model were program (successful POEMB applicants versus unsuccessful applicants), time (in quarters), and time-squared. Time-squared was included so that any curvilinear as well as linear trends could be detected. The interaction terms were program by time and program by time-squared. If receiving a POEMB award affected publication activity, we would expect to see a program by time and/or program by time-squared interaction (e.g., relative differences between groups emerging during the post-POEMB period). The model did indeed show main effects of program, time, and time-squared, and program by time interactions in both the linear and quadratic terms (all terms significant at $p < 0.01$). These effects reflect the difference between groups, the increase over time across all investigators (decelerating in the post-award period), and the greater relative increase of the unsuccessful POEMB applicant group over time.

5.2.2 Journal Title Descriptors

In this section, we discuss findings related to journal article publication activity in terms of (1) longitudinal analysis and (2) "switching" analysis. The journal title descriptors described in Section 5.1.1 are used to characterize the overall disciplinary focus of the articles so that general patterns in the research direction of the individuals in the two POEMB applicant groups can be detected.

5.2.2.1 Longitudinal Analysis

Cardiovascular Journals. Figure 5.2 shows the average number of cardiovascular-designated journal articles over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each time series. The trends indicate increasing numbers for unsuccessful POEMB applicants relative to successful applicants throughout the pre- and post-award periods, consistent with their greater numbers of publications overall.

As with total articles, the repeated measures analysis showed main effects of program, time (both at $p < 0.01$), and time-squared ($p < 0.05$), and an interaction effect of program by time in the linear terms ($p < 0.05$), but no interaction in the quadratic term. These effects reflect the difference between groups, the increase over time across all investigators (decelerating in the post-award period), and the greater relative increase of the unsuccessful POEMB applicant group over time. Correspondingly, simple before-after comparisons showed ten-fold or greater differences between groups for the aggregate, pre-, and post-periods, as well as large differences in gains over time (all significant at $p < 0.01$). Examination of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups. In sum, the difference between POEMB applicant groups on publications in cardiovascular journals was substantial and growing, and notably larger than the difference in total publication activity.

It is possible that POEMB could fail to influence the total number of publications per investigator in a particular journal type (e.g., cardiovascular), and still influence the proportion of publications appearing in that journal type. Therefore, we also examined the ratio of cardiovascular to total publications for those investigators registering one or more publications in **Medline**. This allowed us see changes over time in the mean proportion of articles in cardiovascular journals holding total articles constant. The longitudinal data and their trends are plotted in Figure 5.3. As shown, the unsuccessful POEMB applicants' average ratio is consistently higher than that of the successful POEMB applicants after the first few time periods. Simple before-after comparisons showed large differences between groups for the aggregate (10% to 1%), pre- (9% to 1%), and post- (11% to 1%) periods

Figure 5.2 Mean Number of Publications in Cardiovascular Journals Per Investigator, By Quarter

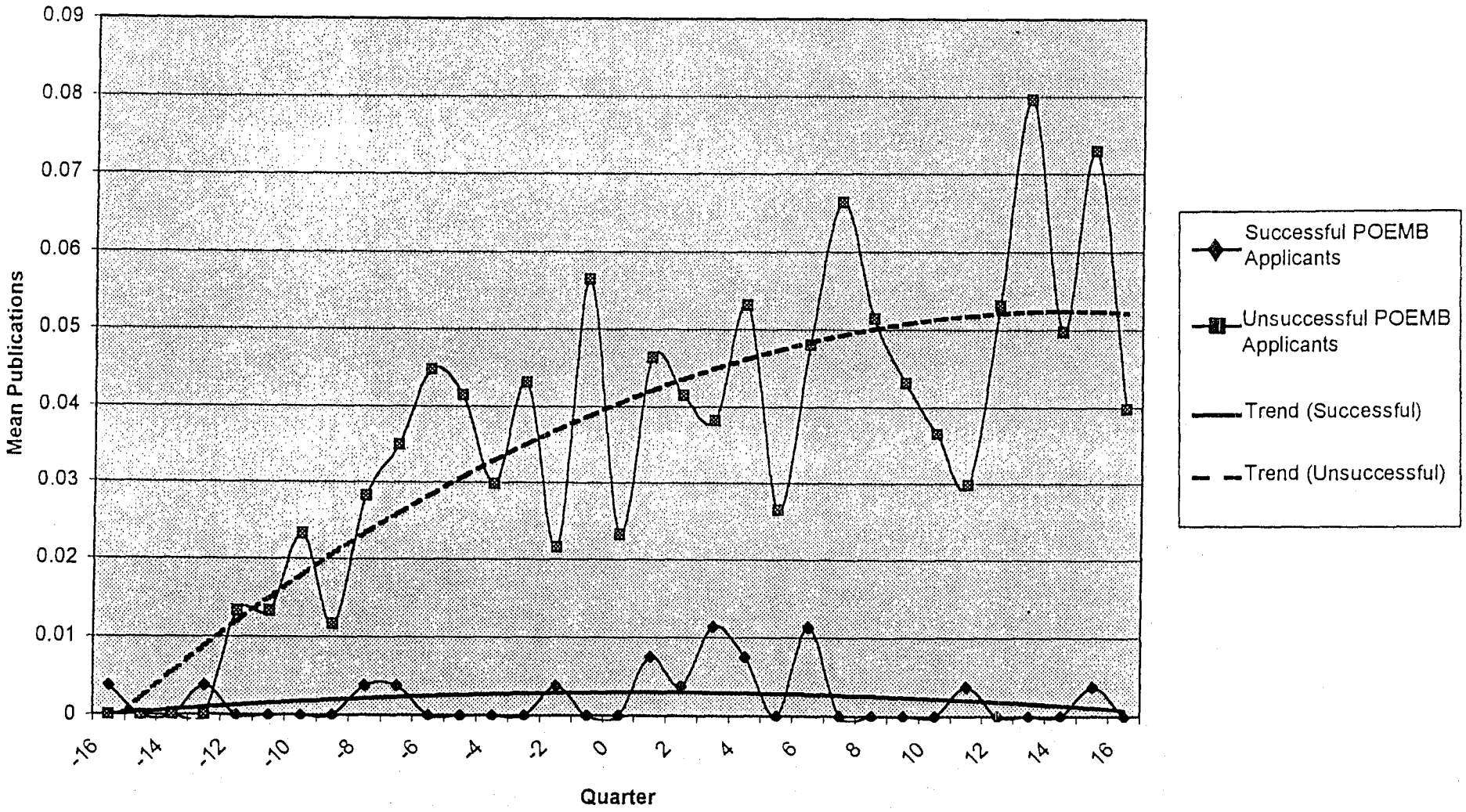
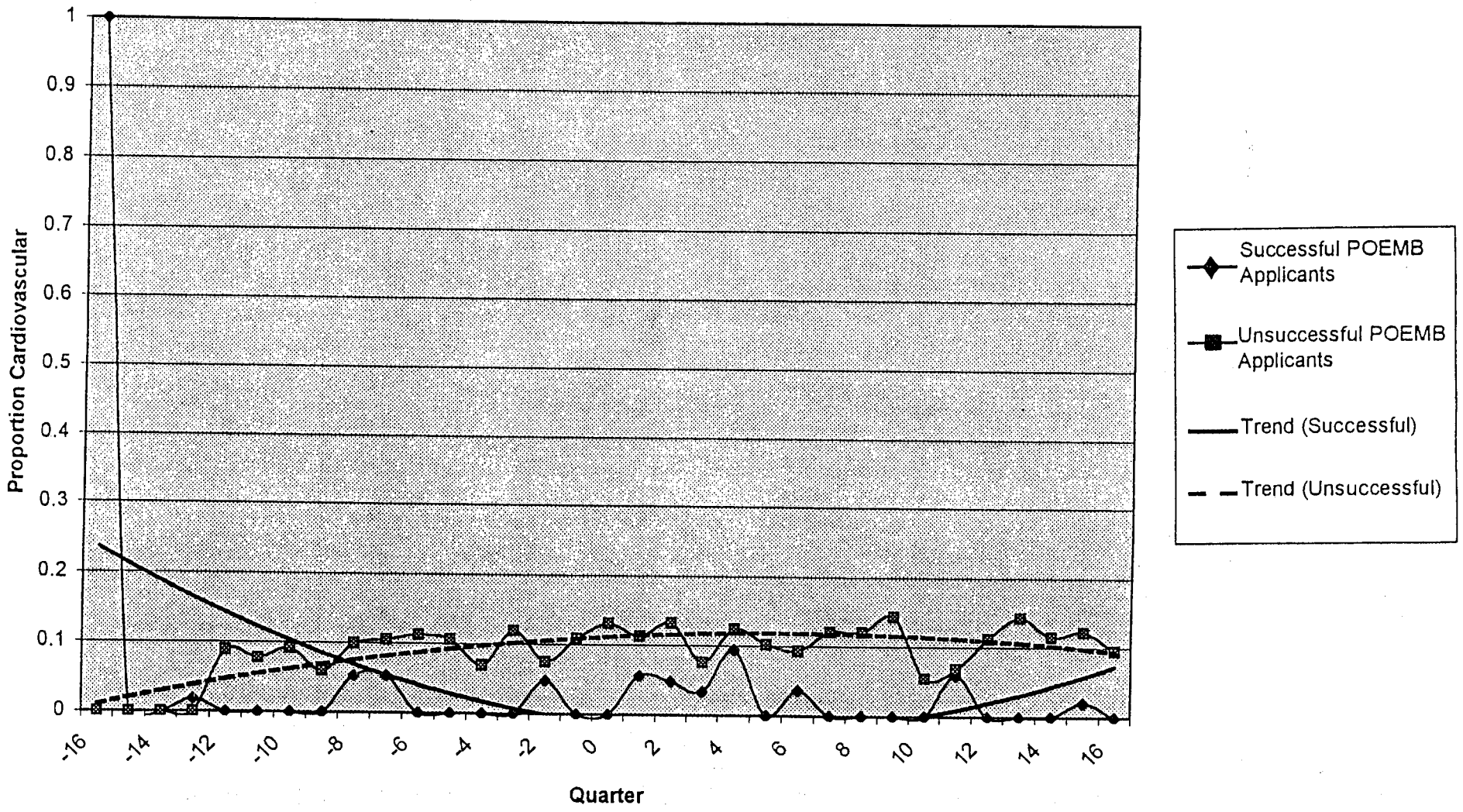


Figure 5.3 Mean Proportion of Total Publications Appearing in Cardiovascular Journals Per Investigator, By Quarter



(all significant at $p < 0.01$).[1] Examination of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups. Thus, there was no evidence from these data that receiving a POEMB grant influenced the numbers or the proportion of publications in cardiovascular journals.

Respiratory Journals. Figure 5.4 shows the average number of respiratory-designated journal articles over time for individuals in the successful and unsuccessful POEMB applicant groups, along with the least squares quadratic trend line through each time series. The trends indicate increasing numbers for unsuccessful POEMB applicants publishing in respiratory journals relative to successful applicants throughout the pre-award period and beginning to decrease late in the post-award period. Given the small numbers and high variability, however, the curvilinear trend among the unsuccessful POEMB applicants should be interpreted with caution.

The repeated measures analysis showed only the main effect of program ($p < 0.01$); it did not show any main effects of time or interaction effects of program by time in the linear or quadratic terms. Similarly, before-after comparisons showed six-fold or greater differences between groups for the aggregate, pre-, and post-periods (all significant at $p < 0.01$), but neither group showed any pre-post gains. Examination of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

As with cardiovascular journals, examination of the ratio of respiratory to total publications lets us see changes over time in the mean proportion of articles in respiratory journals holding total articles constant. The longitudinal data and their trends are plotted in Figure 5.5. Variability is high in both applicant groups, but in most intervals the unsuccessful POEMB applicants' ratio is higher. Simple before-after comparisons showed differences favoring unsuccessful POEMB applicants for the aggregate, pre-award, and post-award periods, though none were statistically significant. Among investigators who published in both periods, the respiratory-to-total ratio increased by a greater amount among successful

---

[1] It should be noted, however, that the individuals contributing scores in the post period are not in all cases the same individuals contributing scores in the pre period. This is because the cardiovascular/total ratio can only be computed for individuals who had at least one Medline-cited publication, and the set of investigators getting publications in the post period is not identical to the set of investigators getting publications in the pre period. Consequently, the ratio data are not strictly longitudinal like the count data. Similarly, a longitudinal analysis could not be conducted on the quarterly data due to high numbers of missing values in each quarter (i.e., cardiovascular/total ratios were only computable in quarters where the investigator submitted one or more applications). This will be the case for all subsequent analyses of rates, ratios, or proportions.

Figure 5.4 Mean Number of Publications in Respiratory Journals Per Investigator, By Quarter

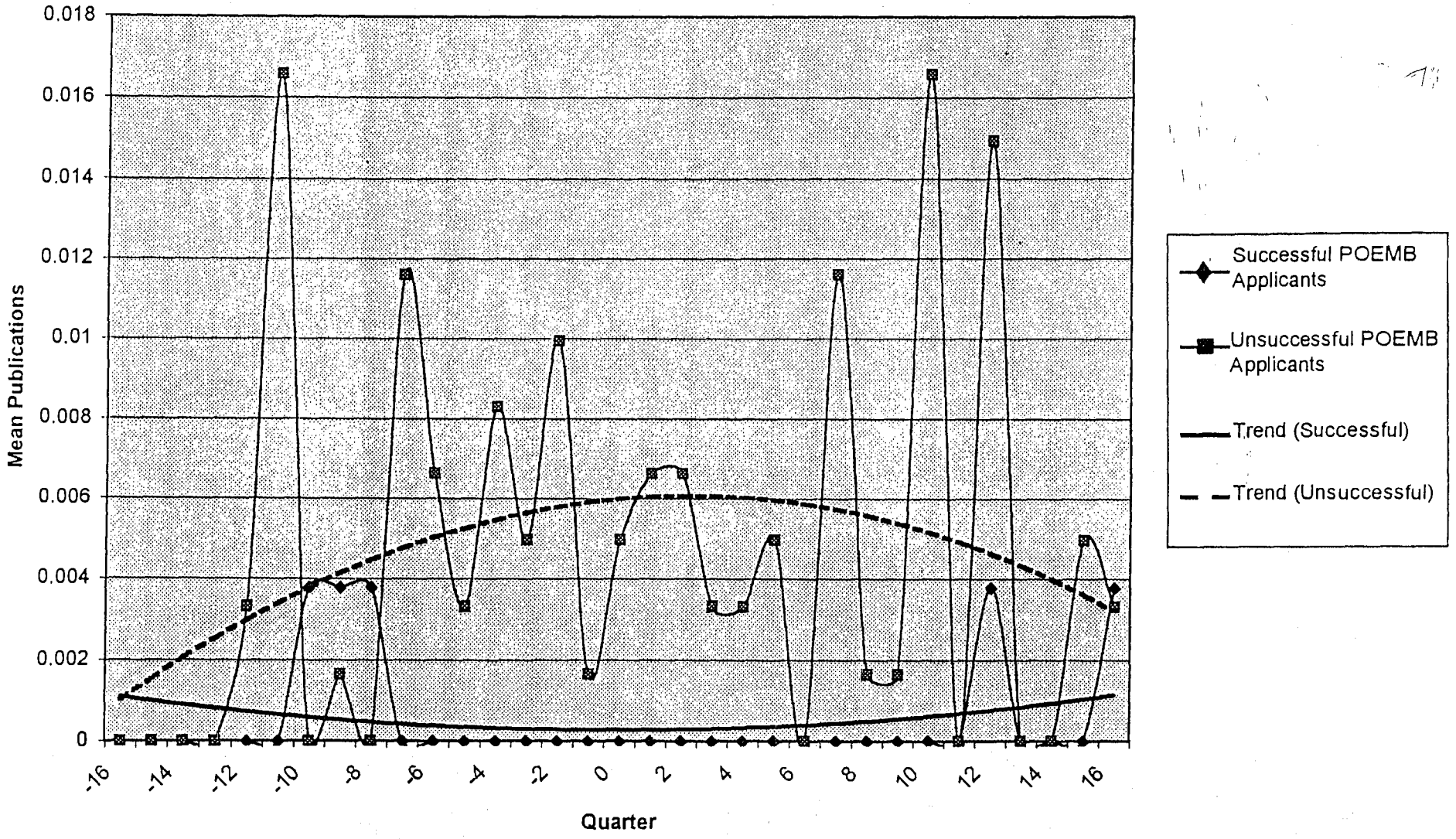
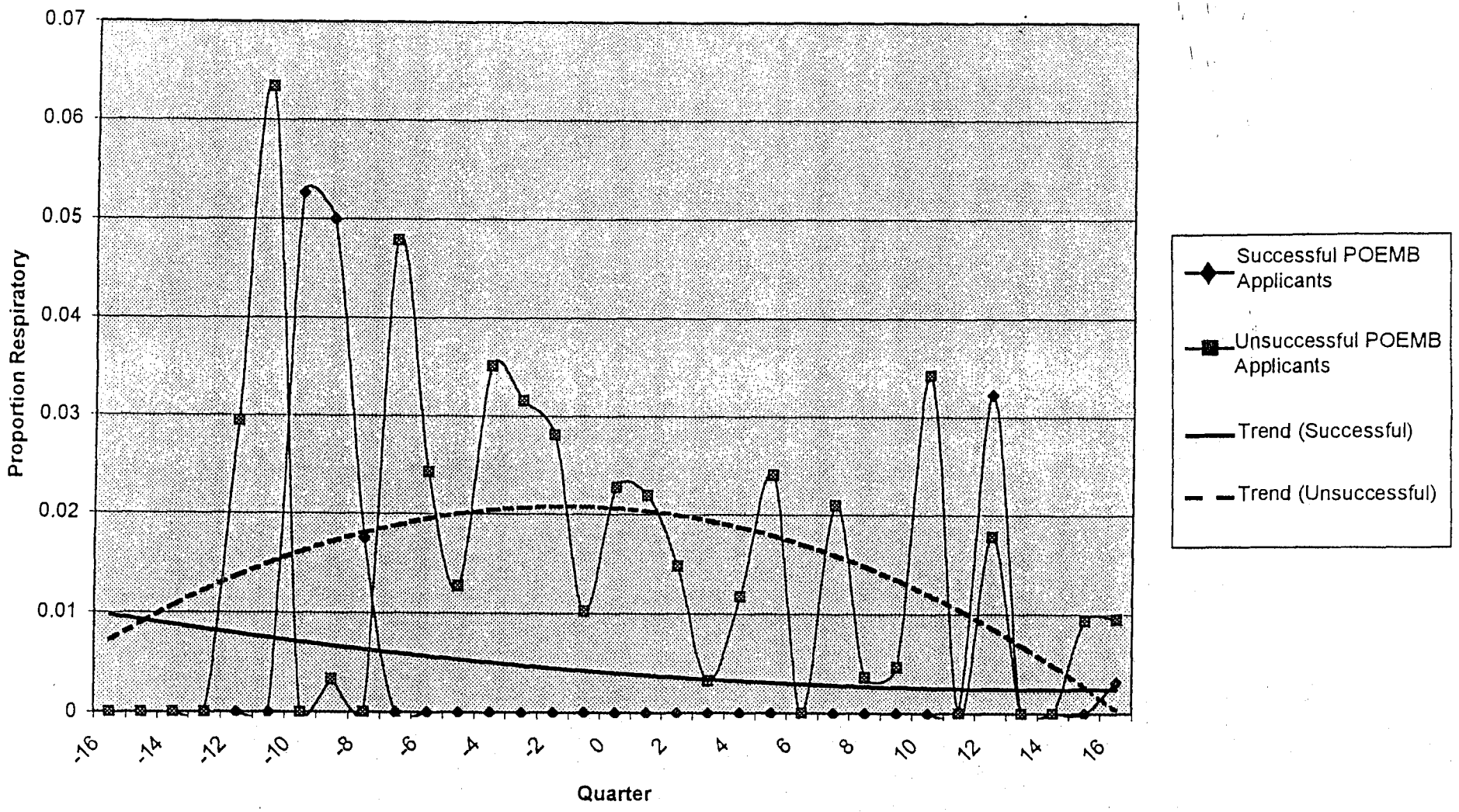


Figure 5.5 Mean Proportion of Total Publications Appearing in Respiratory Journals Per Investigator, By Quarter



POEMB applicants, but the difference was small and nonsignificant. Stratification of this data by status revealed a pattern consistent with the results of the analysis in which all statuses were pooled within the applicant groups. Overall, there was no evidence from these data that receiving a POEMB grant influenced the numbers or the proportion of publications in respiratory journals.

Molecular/Microbiology Journals. Figure 5.6 shows the average number of molecular/microbiology-designated journal articles over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each time series. The trends indicate both groups increasing steadily over time, with the unsuccessful POEMB applicants producing more articles than the successful applicants throughout most of the pre- and post-award periods. The repeated measures analysis showed main effects of program, time, and time-squared (all at $p < 0.01$), no interaction effect of program by time in the linear or term, but an interaction in the quadratic term ($p < 0.05$) consistent with the unsuccessful POEMB applicant group's curvilinearity. Before-after comparisons showed differences between groups for the aggregate, pre-, and post-periods (all significant at $p < 0.01$), but no differences in the pre-post gains, which — as Figure 5.6 shows — were substantial in both groups. Examination of these same data by status showed the POEMB senior investigators with higher mean numbers of publications in molecular/microbiology-designated journals for all periods. Both groups of senior investigators increase from the pre-period to the post-period, with POEMB senior investigators increasing to a greater degree in the post-POEMB period. The total and post-period differences were statistically significant ($p < 0.05$). A very similar pattern is apparent for the trainees in both POEMB applicant groups.

With respect to the ratio of articles in molecular/microbiology journals to total journal output, the longitudinal data and their trends are plotted in Figure 5.7. This pattern differs from those presented previously, in that the successful POEMB applicants began lower than their counterparts, draw even in the latter part of the pre-award period, and pass them in the post-award period. Before-after comparisons show the successful POEMB applicants with a higher ratio over the aggregate period ($p < 0.01$). As expected from the longitudinal plot, the pre-award difference was relatively small and (46% to 42%) nonsignificant, while the post-award difference was substantial (59% to 41%) and statistically significant ($p < 0.01$). This suggests a group by time interaction and is consistent with an explanation that POEMB played a role in the shift in emphasis to molecular/microbiology journals over

Figure 5.6 Mean Number of Publications in Molecular/Microbiology Journals Per Investigator, By Quarter

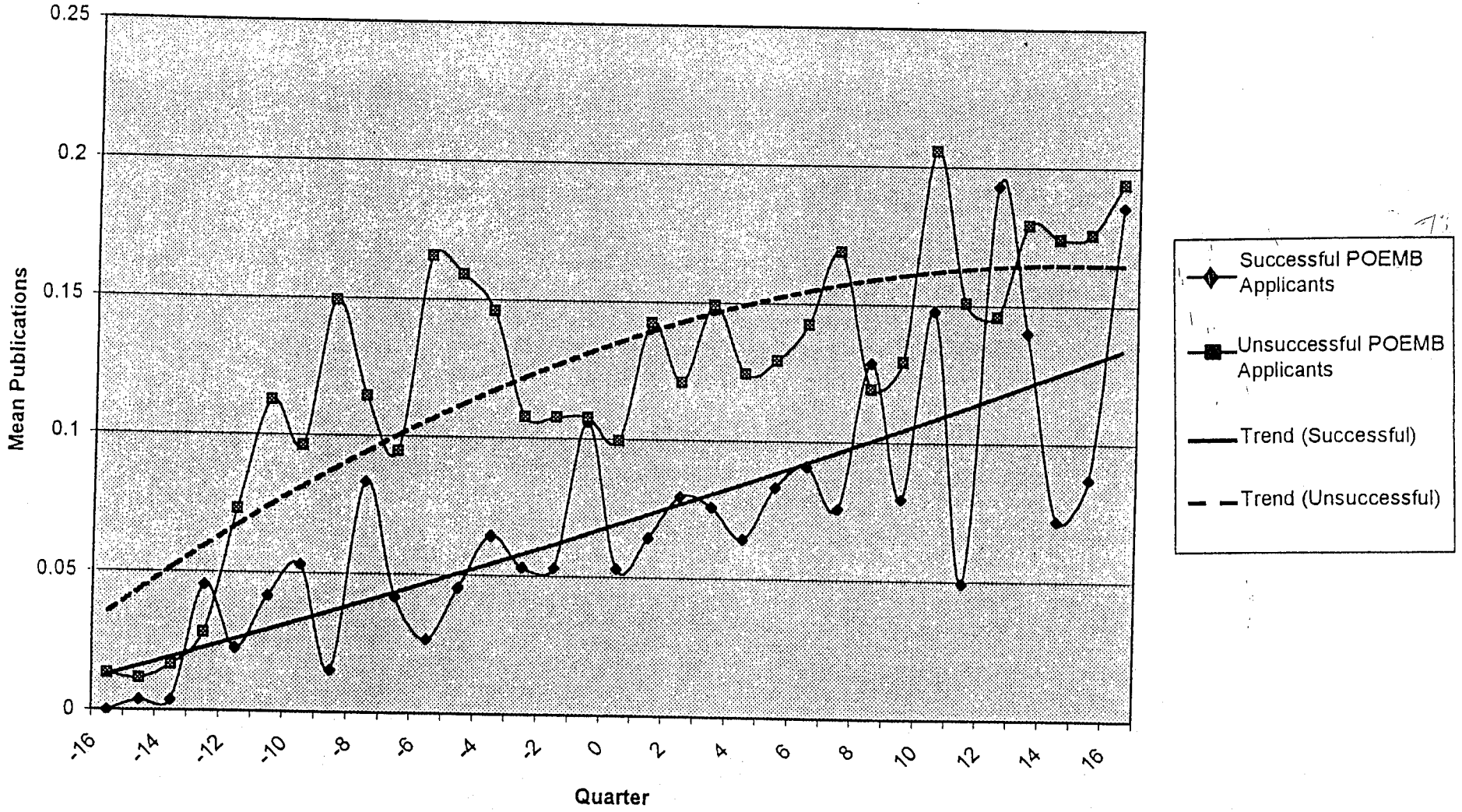
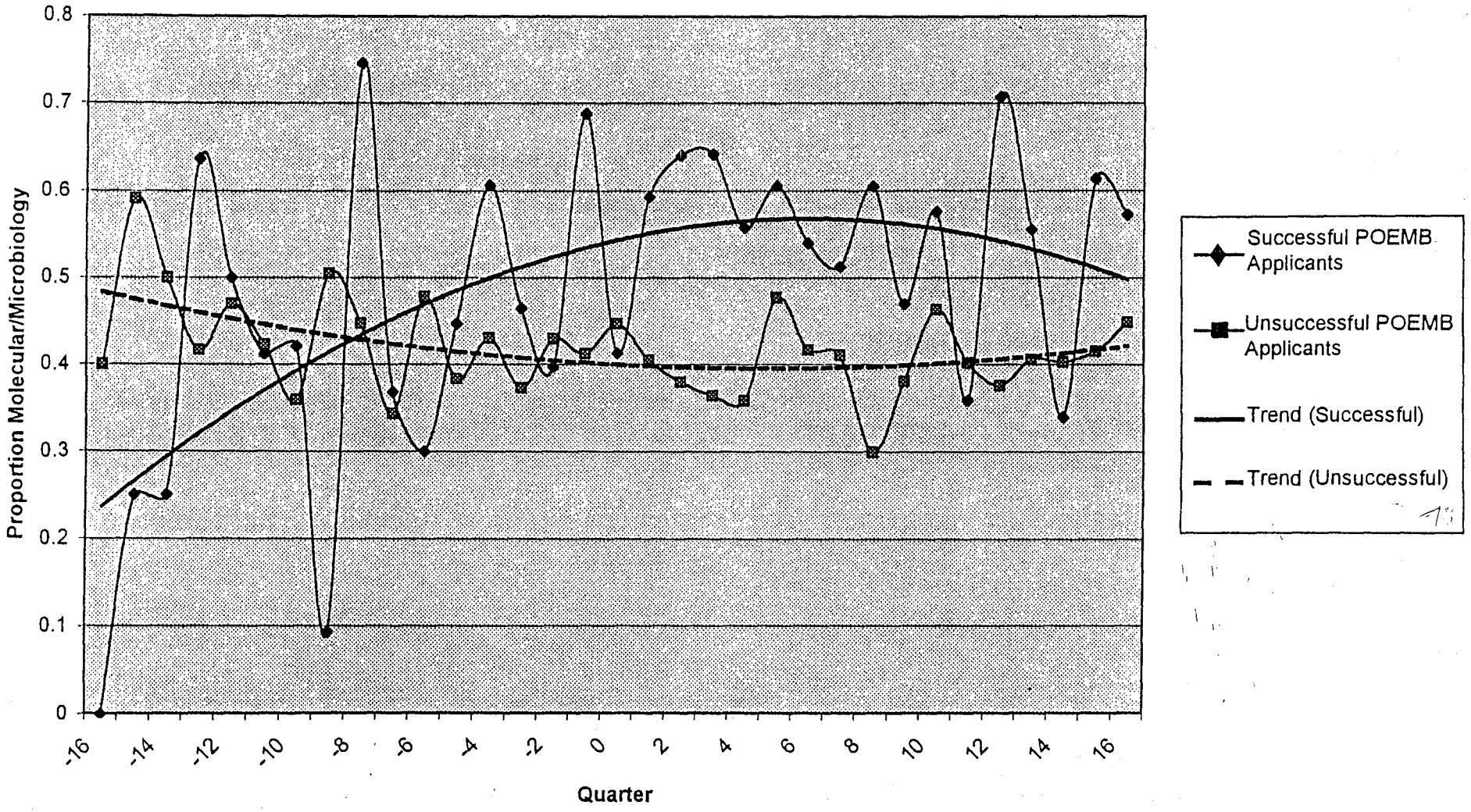


Figure 5.7 Mean Proportion of Total Publications Appearing in Molecular/Microbiology Journals Per Investigator, By Quarter



time. Gain scores also differed (10% to 0%) though the difference was nonsignificant due to reduced *n*'s (only 49 successful POEMB applicants had **Medline** publications in both the pre- and post-award periods). Examination of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.[1]

Cross-Cutting Journals. Figure 5.8 shows the average number of cross-cutting-designated journal articles over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each time series. While there was high variability across quarters in both groups, the trends show unsuccessful POEMB applicants publishing more articles in cross-cutting journals throughout the observation period, with little change in the size of the difference over time. The repeated measures analysis showed main effects of program ($p < 0.01$) and time-squared ($p < 0.05$), but no interaction effect of program by time in the linear or quadratic terms. Before-after comparisons showed differences between groups for the aggregate ($p < 0.01$), pre ($p < 0.05$), and post-periods ($p < 0.01$). The comparison of pre-post gain scores showed that unsuccessful POEMB applicants increased their numbers of publications in cross-cutting journals relative to successful applicants, but the difference was nonsignificant. Stratification of these same data by status revealed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups. A comparison of the total number of articles published by the status 3 and 4 POEMB trainees in journals designated as "cross-cutting" shows that, for the post-POEMB period, the status 4 trainees cited in POEMB progress reports published significantly more articles in these journals than the potential trainees or status 3s ($p < 0.05$). Excluding the status 4s from the analysis would have potentially increased the post-POEMB period difference cited above between the two applicant groups. Thus, including the status 4s in the analysis did not bias the results against the POEMB.

---

[1] A comparison of the ratios of total number of articles published in MB-designated journals by the status 3 and 4 POEMB trainees shows that, for the total and post-POEMB period, the status 3s (potential trainees) published significantly more articles in these journals than the status 4 trainees actually cited in POEMB progress reports ($p < 0.05$). Thus, an "apples-to-apples" comparison of status 3 trainees across applicant groups would show a still stronger association between POEMB and the total number of articles published in MB-designated journals among trainees. However, the fact that the status 3s may have had little or no association with the POEMB grant (judging by the absence of their names from progress reports) muddies interpretation.

Figure 5.8 Mean Number of Publications in Cross-Cutting Journals Per Investigator, By Quarter

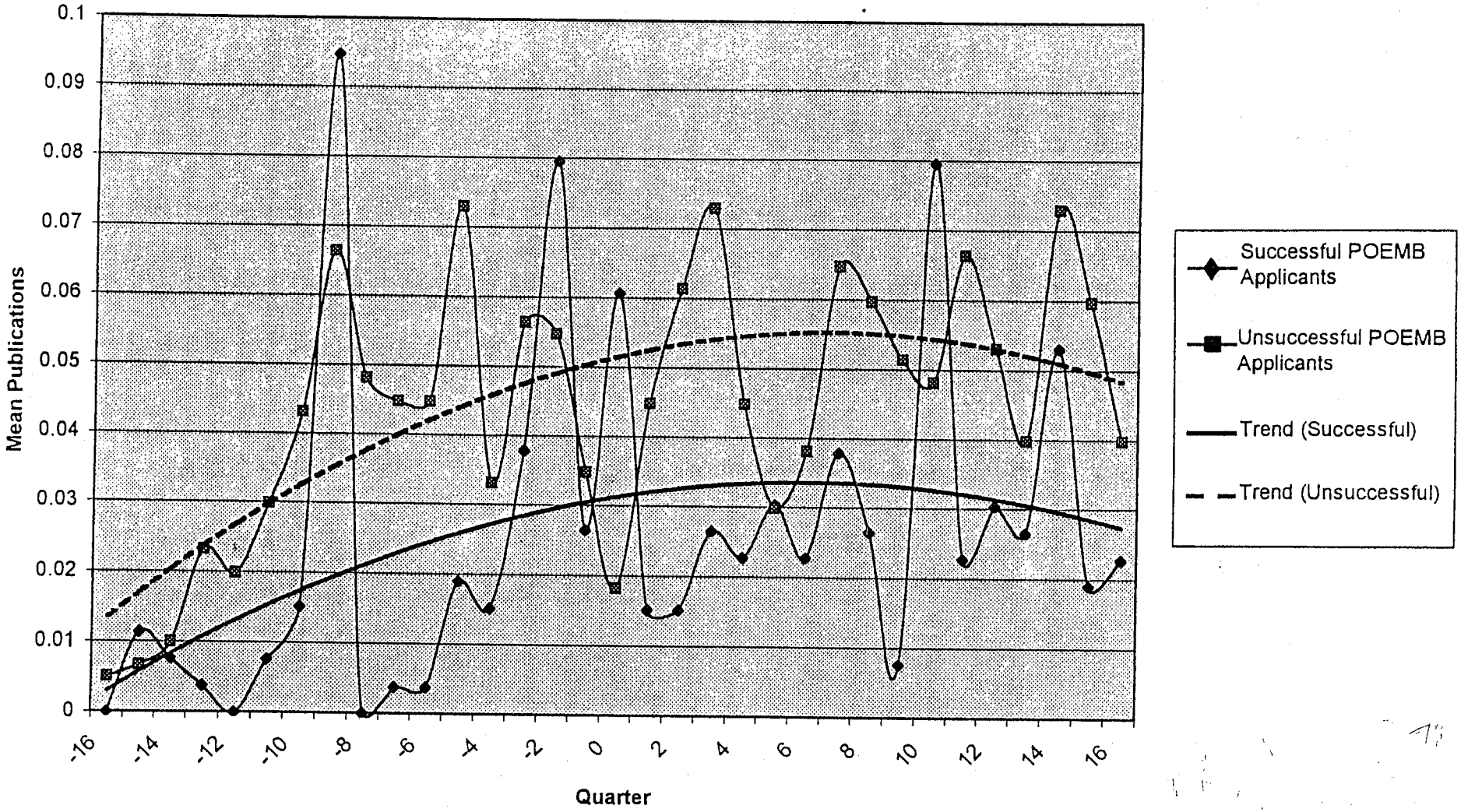


Figure 5.9 shows the ratio of cross-cutting to total publications over time for each group, along with the associated trend lines. These trends are somewhat difficult to interpret due to high variability over time in the successful POEMB applicant group (recall that fewer successful applicants than unsuccessful applicants had publications in any given quarter, so the sample sizes on which the data points are based are lower). Before-after comparisons showed successful applicants with a significantly higher mean ratio during the pre-award period (26% to 16%, $p<0.05$), which disappeared and changed direction in the post-award period (11% to 14%, not statistically significant). Examination of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups. Thus, receiving a POEMB grant was associated with a decrease in the mean proportion of articles appearing in cross-cutting journals. A comparison of the status 3 and 4 POEMB trainees' ratios of articles in cross-cutting journals to total publications shows that, for the total period, the status 4 trainees cited in POEMB progress reports exhibit a significantly higher ratio ($p<0.05$) than the status 3s (or potential trainees). Thus, including the status 4s in the analysis did not bias the results against the POEMB.

"Other" Journals. Figure 5.10 shows the average number of "other"-designated journal articles over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each time series. As shown, the two series begin at approximately the same level, after which the unsuccessful applicants' series increases steadily while the successful applicants' series remains flat until late in the post-award series. The repeated measures analysis showed main effects of program, time, and time-squared, and interaction effects of program by time in both the linear and quadratic terms (all significant at $p<0.01$). Correspondingly, before-after comparisons showed two-fold or greater differences between groups for the aggregate, pre-, and post-periods, as well as still larger relative differences (six-fold) in gains over time (all significant at $p<0.01$). Stratification of these same data by status revealed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

Figure 5.11 shows the "other" to total ratio over time for each group. While there are definite fluctuations (primarily among successful POEMB applicants), the overall trends indicate little change over time and little difference between groups. Correspondingly, the before-after comparisons showed no significant differences over the aggregate, pre-award, or post-award periods, and no differences in

Figure 5.9 Mean Proportion of Total Publications Appearing in Cross-Cutting Journals Per Investigator, By Quarter

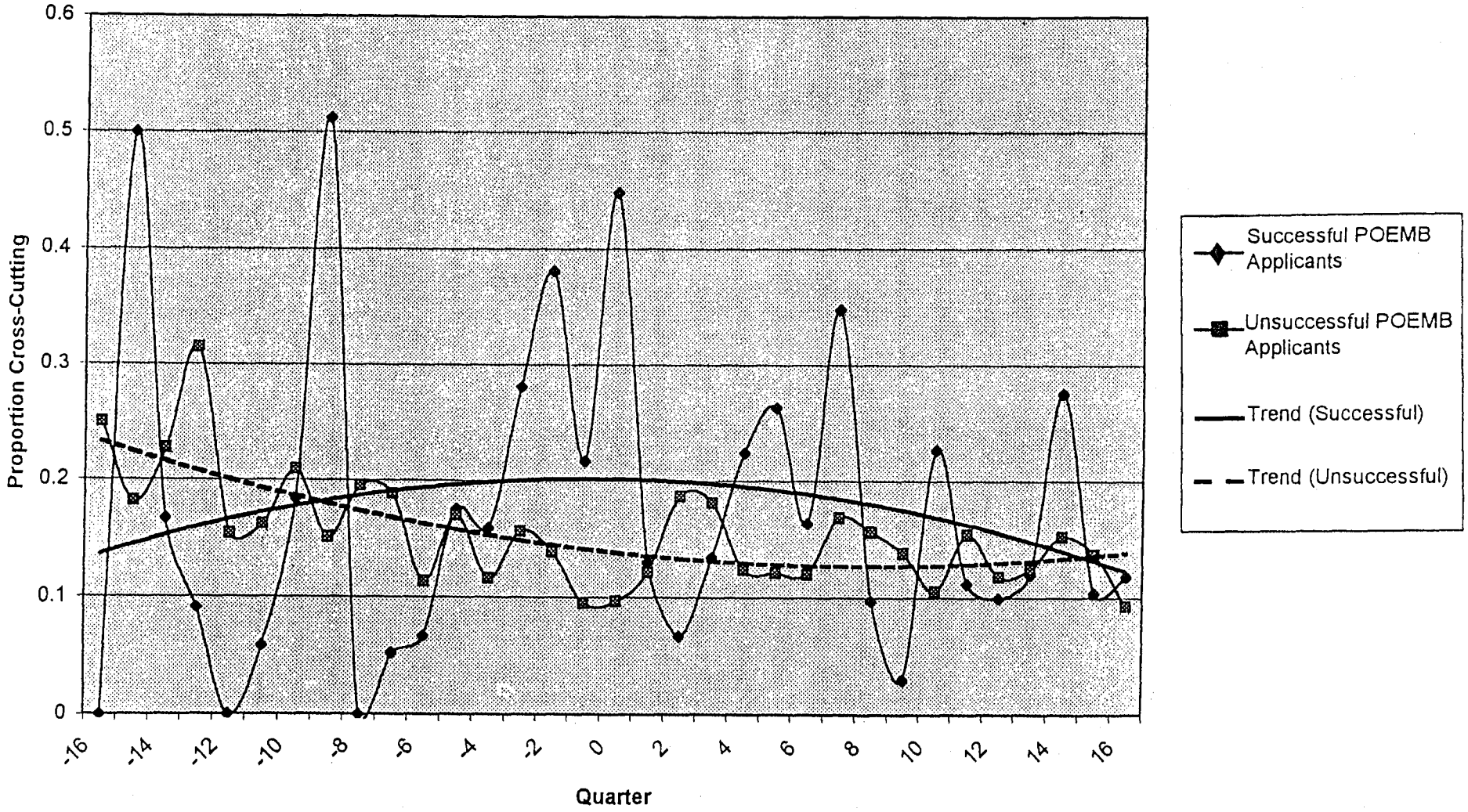


Figure 5.10 Mean Number of Publications in "Other" Journals Per Investigator, By Quarter

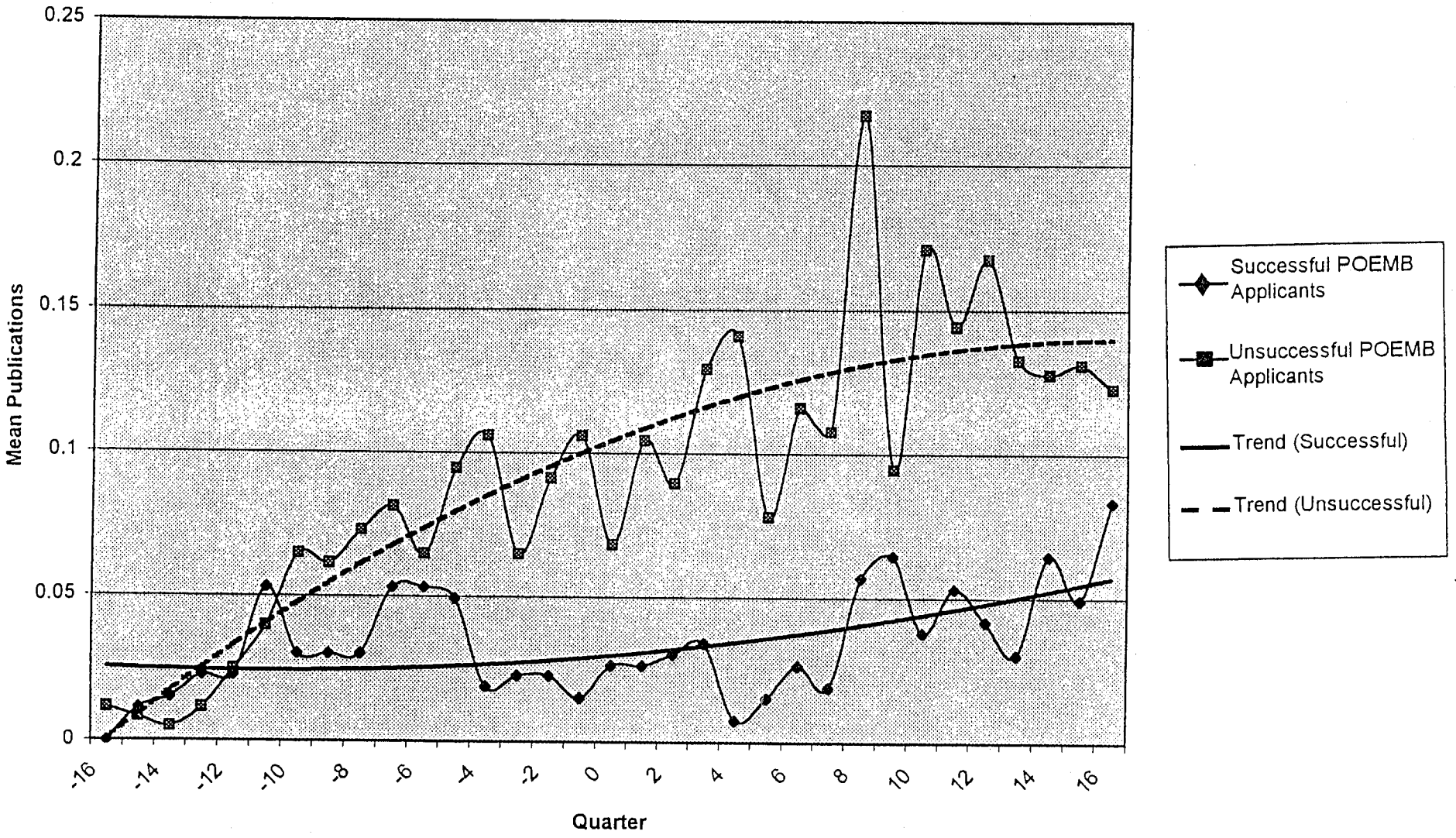
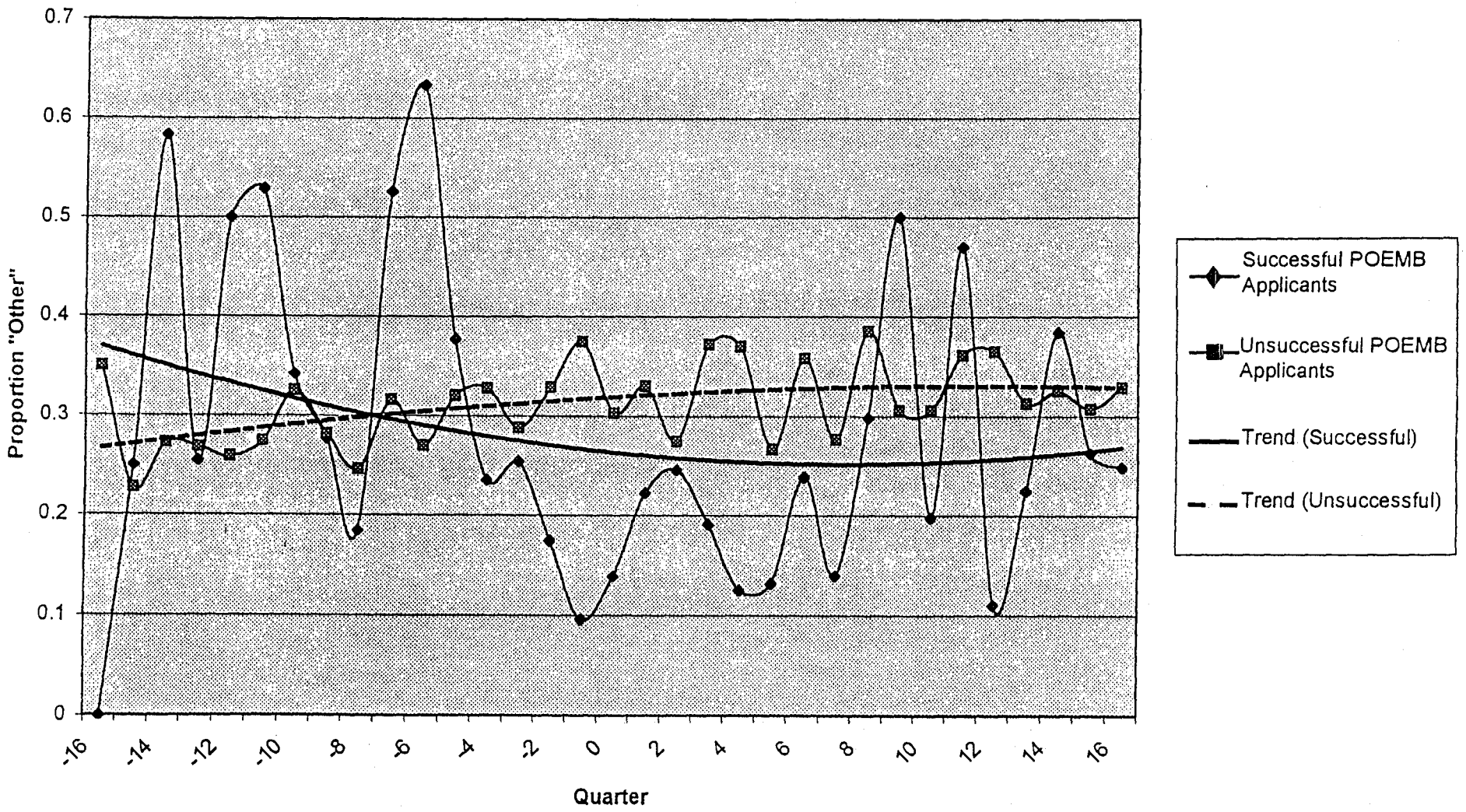


Figure 5.11 Mean Proportion of Total Publications Appearing in "Other" Journals Per Investigator, By Quarter



gain scores. Examination of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups. Thus, there is no evidence from these data that receiving a POEMB grant influenced the numbers or the proportion of publications in "other" journals.

5.2.2.2 Analysis of "Switching"

Tables 5.3 and 5.4 show the complete transition matrices for successful and unsuccessful POEMB applicants, respectively. The transition matrices show the probability of all possible journal title category transitions from pre-POEMB to post-POEMB status. There are eight possible pre- and post-states:

- None - no **Medline** publications
- HL only – publication in cardiovascular and/or respiratory journals only
- MB only - publication in molecular/microbiology journals only
- HL/MB - publication in HL/MB or (HL and MB) journals
- Cross-cutting or "other" only - publication in cross-cutting or "other" journals only
- HL/other - publication in HL and (cross-cutting or "other") journals
- MB/other - publication in MB and (cross-cutting or "other") journals
- HL/MB/other - publication in HL/MB or (HL and MB awards) and (cross-cutting or "other") journals

Each pre-POEMB state has a separate set of transition probabilities that sum to 100%. In Table 5.3, for example, among successful POEMB applicants who had no **Medline** publications pre-award, 68% had no publications post-award, 15% had only MB publications, 8% had only cross-cutting or "other" publications, 7% had MB and cross-cutting or "other" publications, etc.

The darkly shaded cells represent the 16 transition states that could be categorized as "switching," so designated because each represents a transition from a non-HL to HL state or a non-MB to MB state. To cover all transitions of potential interest, we characterized switching under two definitions: a more restrictive and a less restrictive. The more restrictive or "pure" switch conditions are those in which the investigator transitions from an MB to an HL state or vice versa. These are:

- HL to MB. This includes HL only to MB only, HL only to MB/other, HL/other to MB only, HL/other to MB/other.

Table 5.3 Pre-POEMB Award to Post-POEMB Award Journal Title Category Transition Matrix for Articles Published by Successful POEMB Applicants

Pre-POEMB Application	Post-POEMB Application								
Frequency Row Percent	HL	HL/ MB	HL/MB /"other"	HL/ "other"	MB	MB/ "other"	none	"other"	Total
HL	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0
HL/MB	0 0%	0 0%	0 0%	0 0%	1 100 %	0 0%	0 0%	0 0%	1
HL/MB/ "other"	0 0%	0 0%	3 75%	0 0%	1 25%	0 0%	0 0%	0 0%	4
HL/ "other"	0 0%	0 0%	0 0%	1 50%	1 50%	0 0%	0 0%	0 0%	2
MB	0 0%	0 0%	0 0%	0 0%	5 56%	2 22%	1 11%	1 11%	9
MB/ "other"	0 0%	0 0%	6 18%	0 0%	6 18%	17 50%	4 12%	1 3%	34
none	0 0%	1 0.5%	2 1%	0 0%	30 15%	15 7%	139 68%	16 8%	203
other	1 9%	0 0%	0 0%	0 0%	0 0%	3 27%	7 63%	0 0%	11
Total	1	1	11	1	44	37	151	18	264

Table 5.4 Pre-POEMB Award to Post-POEMB Award Journal Title Category Transition Matrix for Articles Published by Unsuccessful POEMB Applicants

Pre-POEMB Application	Post-POEMB Application								
Frequency Row Percent	HL	HL/ MB	HL/MB / "other"	HL/ "other "	MB	MB/ "other"	none	other	Total
HL	2 14%	0 0%	1 7%	8 57%	1 7%	0 0%	1 7%	1 7%	14
HL/MB	0 0%	0 0%	1 100%	0 0%	0 0%	0 0%	0 0%	0 0%	1
HL/MB/ "other"	0 0%	0 0%	18 53%	5 15%	3 9%	8 24%	0 0%	0 0%	34
HL/ "other"	4 11%	0 0%	12 34%	15 43%	0 0%	1 3%	1 3%	2 6%	35
MB	1 2%	0 0%	3 5%	2 3%	20 32%	24 39%	11 18%	1 2%	62
MB/"other"	0 0%	2 2%	21 20%	5 5%	4 4%	56 54%	9 9%	6 6%	103
none	7 2%	4 1%	7 2%	11 4%	36 12%	35 12%	151 52%	42 14%	293
other	2 3%	1 2%	6 10%	3 5%	5 8%	22 36%	10 16%	12 20%	61
Total	16	7	69	49	69	146	183	64	603

- HL to HL/MB. This includes HL only to HL/MB, HL only to HL/MB/other, HL/other to HL/MB, and HL/other to HL/MB/other.
- MB to HL. This includes MB only to HL only, MB only to HL/other, MB/other to HL only, MB/other to HL/other.
- MB to HL/MB. This includes MB only to HL/MB, MB only to HL/MB/other, MB/other to HL/MB, MB/other to HL/MB/other.

As shown in Tables 5.3 and 5.4, a relatively small number of investigators fell into any of the 4 "pure" switch transition states, and most of these were in the unsuccessful POEMB applicant group. Specifically, 3 were in the HL to MB group (2 unsuccessful applicants, 1 successful applicant), 13 were in the HL to HL/MB group (all unsuccessful applicants), 8 were in the MB to HL group (all unsuccessful applicants), and 32 were in the MB to HL/MB group (26 unsuccessful applicants, 6 successful applicants). In percent terms, the MB to HL/MB difference between groups was two-fold (4% to 2%), and not statistically significant.

The lightly shaded cells in the tables represent the additional less restrictive switch conditions. These were defined as follows:

- **No HL to some HL.** This includes none to HL only, none to HL/MB, none to HL/other, none to HL/MB/other, MB only to HL only, MB only to HL/MB, MB only to HL/other, MB only to HL/MB/other, other to HL only, other to HL/MB, other to HL/other, other to HL/MB/other, MB/other to HL only, MB/other to HL/MB, MB/other to HL/other, MB/other to HL/MB/other.
- **No MB to some MB.** This includes none to MB only, none to HL/MB, none to MB/other, none to HL/MB/other, HL only to MB only, HL only to HL/MB, HL only to MB/other, HL only to HL/MB/other, other to MB only, other to HL/MB, other to MB/other, other to HL/MB/other, HL/other to MB only, HL/other to HL/MB, HL/other to MB/other, HL/other to HL/MB/other.
- **No HL/MB to some HL/MB.** This includes none to HL/MB, none to HL/MB/other, HL only to HL/MB, HL only to HL/MB/other, MB only to HL/MB, MB only to HL/MB/other, other to HL/MB, other to HL/MB/other, HL/other to HL/MB, HL/other to HL/MB/other, MB/other to HL/MB, MB/other to HL/MB/other.

These switching categories are less restrictive in that investigators with and without pre-POEMB publication activity are grouped together. As with grant applications and awards in Chapter 4, it is

arguable that those who have no publications in the pre-POEMB period are not truly switching when they publish in the post-POEMB period; rather they are simply beginning to get published in their chosen field. They are worth examining, however, particularly since the relatively small number of successful POEMB applicants with **Medline**-cited publications in the pre-POEMB period (61) greatly limits their potential to demonstrate switching on this indicator.

Of the 867 investigators, 77 met the No HL to some HL condition. Unsuccessful POEMB applicants were more than three times as likely to fall into this group as were successful applicants (11% to 3%), and this difference was statistically significant ($p<0.01$). Among the subset of investigators who began in the No HL condition (and therefore had the opportunity to switch), the relative frequencies of the two groups showed a similar pattern (13% and 4% for unsuccessful and successful applicants, respectively).

More investigators (177) met the No MB to some MB condition. This time, unsuccessful POEMB applicants only slightly outpaced their counterparts (21 to 19%, n.s.). Among the subset of investigators who began in the No MB condition, however, the difference was somewhat greater (31% to 24%).

As expected, a smaller number (58) met the No HL/MB to some HL/MB condition. Unsuccessful POEMB applicants were more likely to fall into this group (8% to 3%, $p<0.01$). Among the subset of investigators who began in the No HL/MB condition, the relative frequencies of the two groups showed a similar pattern (6% and 3% for unsuccessful and successful applicants, respectively).

In sum, there were several switch patterns in which unsuccessful POEMB applicants transitioned at higher rates than did successful applicants, others in which there was no significant difference, and none in which successful POEMB applicants transitioned at higher rates than did unsuccessful applicants. Thus, these data offer no evidence that receiving a POEMB grant increased the probability of investigator switches of the sort hypothesized, at least as measured by the disciplinary focus of the journals in which they published.

5.2.3 Article Descriptors - MeSH Terms

As in the previous section, we discuss findings related to journal article publication activity in terms of (1) longitudinal analysis and (2) "switching" analysis. However, in this section, the MeSH term categories described in Section 5.1.1 are used to characterize the overall focus and/or subject matter of the

articles in order to detect the general patterns in the research direction of the individuals in the two POEMB applicant groups.

5.2.3.1 Longitudinal Analysis

Heart/Lung Terms. Figure 5.12 shows the average number of articles with heart/lung-related MeSH terms over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each time series. The data show articles with heart/lung MeSH terms beginning at similar levels in both groups, then diverging about midway through the pre-award observation period (with unsuccessful applicants increasing faster), and continuing to diverge throughout much of the post-award period. The repeated measures analysis showed main effects of program, time (both at $p < 0.01$), and time-squared ($p < 0.05$), and an interaction effect of program by time in the linear term ($p < 0.01$), but no interaction in the quadratic term. Simple before-after comparisons showed differences between groups for the aggregate, pre, and post periods, as well as large differences in gains over time (all significant at $p < 0.01$). Stratification of these same data by status revealed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups. A comparison of the average number of articles with heart/lung-related MeSH terms published by the status 3 and 4 POEMB trainees shows that, for the total period, the status 4 trainees cited in POEMB progress reports exhibit a significantly higher ratio ($p < 0.05$) than the status 3s (or potential trainees). Thus, excluding the status 4s from the analysis may have increased the between-applicant group difference slightly, but not by a significant amount.

It is possible that POEMB could fail to influence the total number of articles generating heart/lung-related MeSH terms and still influence the proportion of articles published that generated these terms. Therefore, we also examined the ratio of articles generating heart/lung-designated MeSH terms to total publications for those investigators registering one or more publications in **Medline**. This allowed us to see changes over time in the mean proportion of articles generating heart/lung-designated (and other) MeSH terms holding total articles constant. The longitudinal data and their trends are plotted in Figure 5.13. While high variability in the successful applicants' series hinders interpretability somewhat, the crossing of the trend lines suggests a group by time interaction, in which the unsuccessful applicants' ratio is below that of the successful applicants' through much of the pre-award period, but

Figure 5.12 Mean Number of Journal Publications With Heart/Lung MeSH Terms Per Investigator, By Quarter

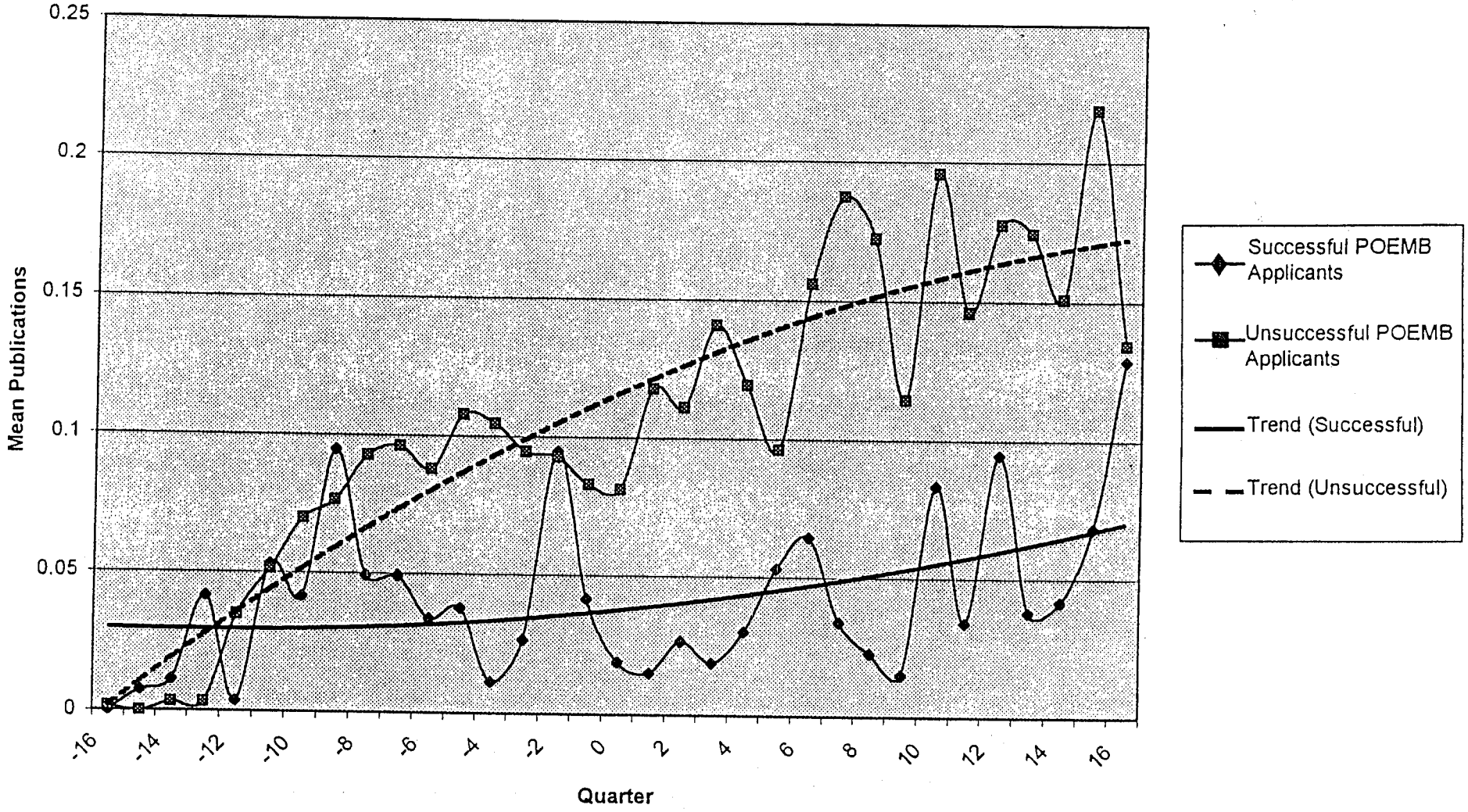
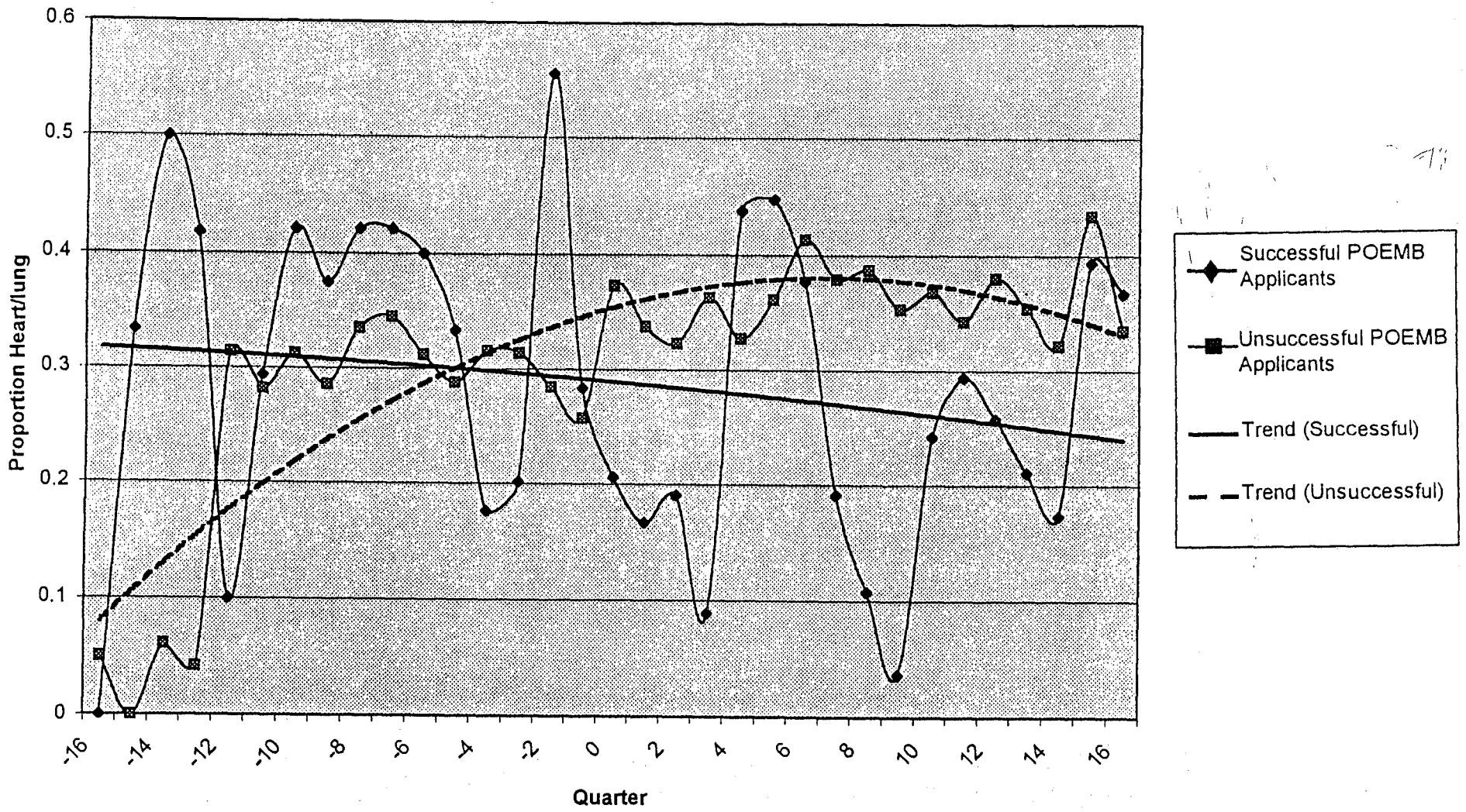


Figure 5.13 Mean Proportion of Journal Publications With Heart/Lung MeSH Terms Per Investigator, By Quarter



higher thereafter.[1] Before-after comparisons revealed significantly higher proportions among unsuccessful POEMB applicants over the aggregate period ($p < 0.05$). There was no difference over the pre-award period, but a difference over the post-award period (37% to 26%, $p < 0.05$), consistent with an interaction effect. The difference in gain scores was not significant, in large part due to low power (only 49 successful POEMB applicants had publications in both the pre- and post-periods, a requirement for this contrast). Stratification by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

In addition to examining the number of articles with heart/lung-related MeSH terms, we also looked at the number of articles with *only* heart/lung MeSH terms (i.e., no co-occurring molecular biology-related MeSH terms). The longitudinal data and their associated trends are shown in Fig. 5.14. Here the difference is even more pronounced, with successful POEMB applicants showing virtually no upward trend. The repeated measures analysis showed main effects of program, time, and time-squared (all at $p < 0.01$), and interaction effects of program by time in the linear ($p < 0.01$) and quadratic ($p < 0.05$) terms. Simple before-after comparisons showed differences between groups for the aggregate, pre-, and post-periods, with the size of the difference more than doubling over the pre-post periods (all significant at $p < 0.01$). Examination of these same data by status revealed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

Figure 5.15 shows the ratio of articles with only heart/lung MeSH terms to total articles over time for each group. The trend lines suggest an even more pronounced interaction effect than when co-occurring terms were included. Before-after comparisons revealed significantly higher proportions among unsuccessful POEMB applicants for the aggregate, pre-, and post-award periods, as well as a significant difference in gain scores ($p < 0.01$), consistent with an interaction. Before-after comparisons revealed significantly higher proportions among unsuccessful POEMB applicants over the aggregate period ($p < 0.01$). There was no difference over the pre-award period, but a difference over the post-award period (19% to 10%, $p < 0.01$), consistent with an interaction effect. Examination of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

---

[1] As with journal publications, the ratio data are not strictly longitudinal like the count data, and a longitudinal analysis could not be conducted on the quarterly data due to high numbers of missing values in each quarter. This will be the case for all subsequent analyses of rates, ratios, or proportions.

Figure 5.14 Mean Number of Journal Publications With Heart/Lung-Only MeSH Terms Per Investigator, By Quarter

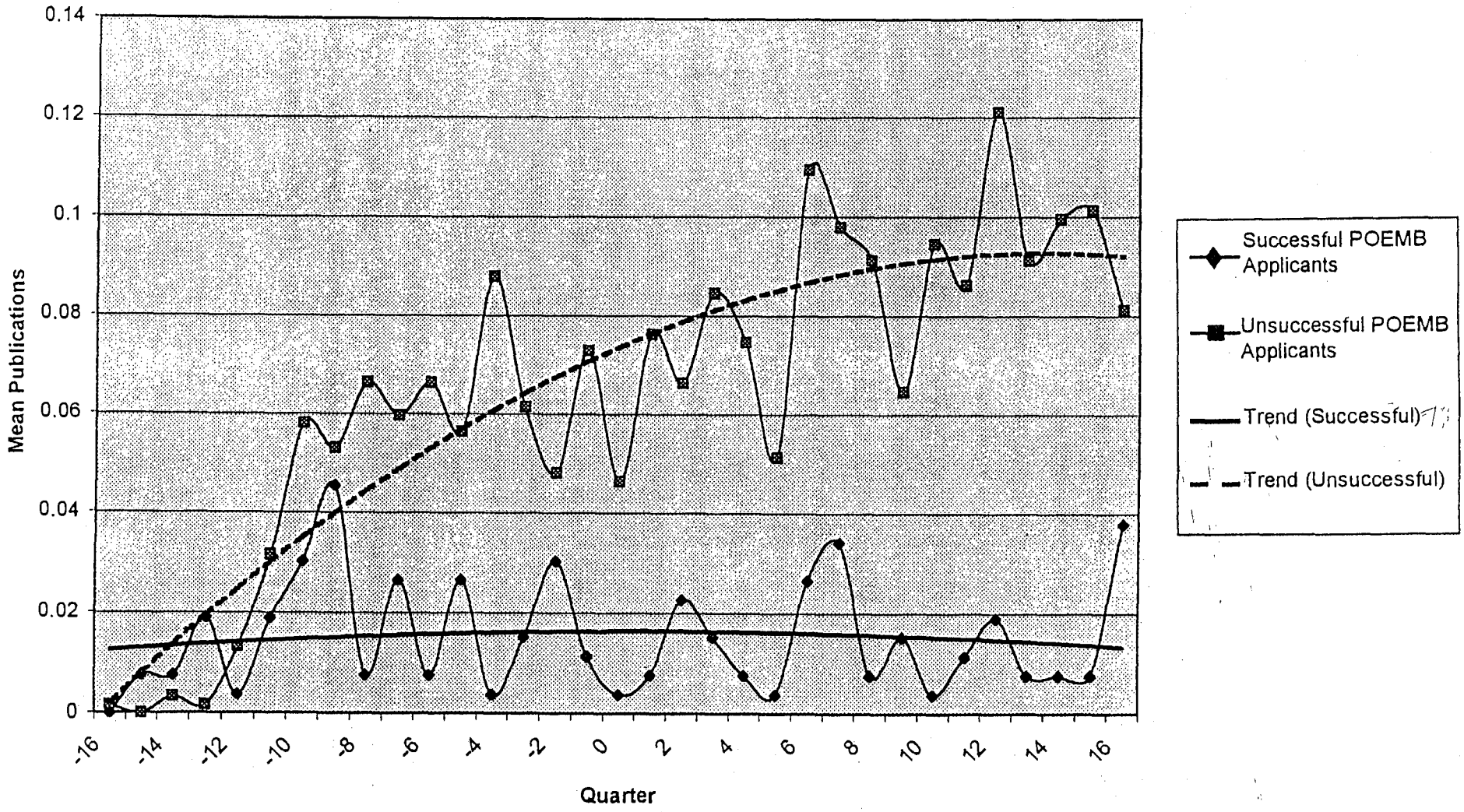
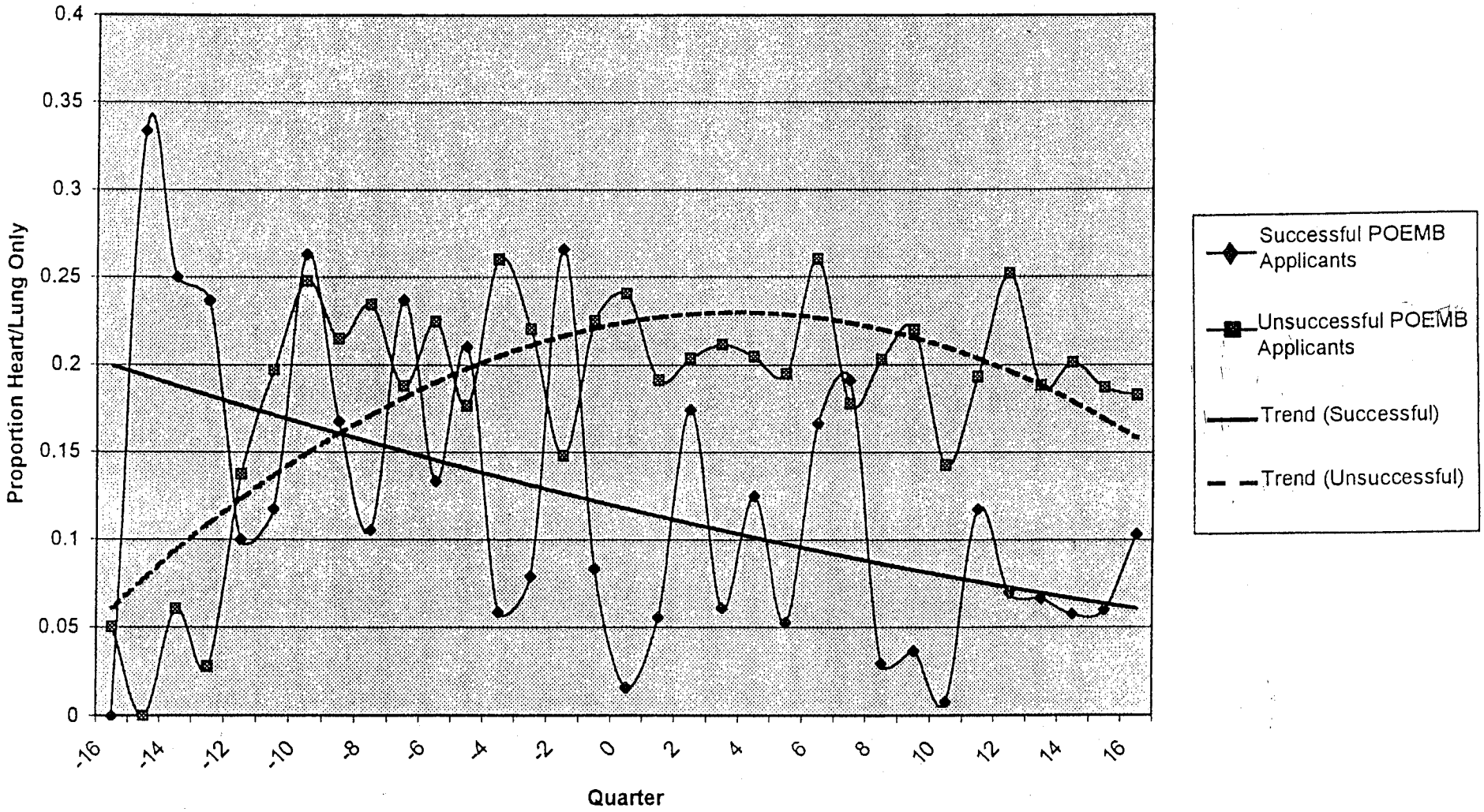


Figure 5.15 Mean Proportion of Journal Publications With Heart/Lung Only MeSH Terms Per Investigator, By Quarter



Molecular Biology Terms. Figure 5.16 shows the average number of articles with molecular biology-related MeSH terms over time for successful and unsuccessful POEMB applicants, along with the associated trend lines. As shown, both groups increased over time on this dimension, with the unsuccessful POEMB applicants increasing at a somewhat faster pace. The repeated measures analysis showed main effects of program and time (both at $p < 0.01$) and an interaction effect of program by time in the linear ($p < 0.01$) and quadratic ($p < 0.05$) terms. Simple before-after comparisons showed differences between groups for the aggregate, pre-, and post-award periods (all significant at $p < 0.01$), with a smaller yet still significant ($p < 0.05$) difference in gain scores. Stratification of these same data by status revealed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

Figure 5.17 shows the ratio of articles with molecular biology MeSH terms to total articles over time for each applicant group. As shown, the ratio is higher among successful POEMB applicants for most of the pre-award period and all of the post-award period. Note also that the proportions in both groups are 0.5 or higher in almost every period, indicating that the majority of articles published by these investigators during the observation period generated molecular biology MeSH terms. Before-after comparisons revealed that successful POEMB applicants' average proportion was significantly higher over the aggregate period ($p < 0.01$). As suggested by the figure, the pre-award difference (66% to 56%, $p < 0.05$) increased in the post-award period (82% to 65%, $p < 0.01$). Stratification of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

As with heart/lung MeSH terms, we also looked at the number of articles with *only* molecular biology MeSH terms (i.e., no co-occurring heart/lung terms). As shown in Figure 5.18, successful POEMB applicants increased at a slightly more comparable rate than when articles with co-occurring heart/lung terms were lumped in. The repeated measures analysis showed main effects of program, time (both at $p < 0.01$), and time-squared ($p < 0.05$), and an interaction effect of program by time in the linear term ($p < 0.05$), but no interaction in the quadratic term. The before-after comparisons still showed differences between groups for the aggregate, pre-, and post-award periods (all significant at $p < 0.01$), but no differences in the pre-post gain scores. Examination of these same data by status revealed

Figure 5.16 Mean Number of Journal Publications With Molecular Biology MeSH Terms Per Investigator, By Quarter

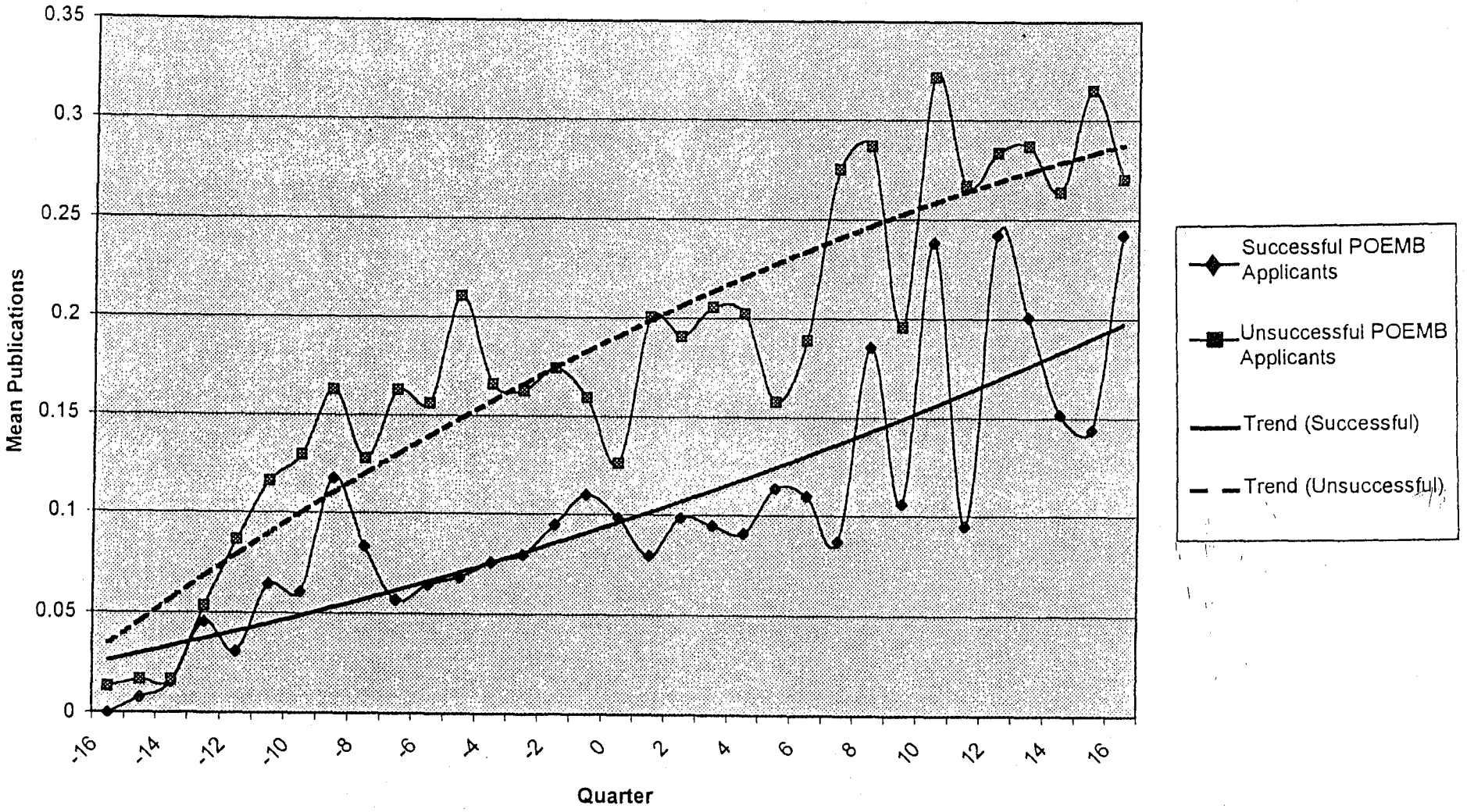


Figure 5.17 Mean Proportion of Journal Publications With Molecular Biology MeSH Terms Per Investigator, By Quarter

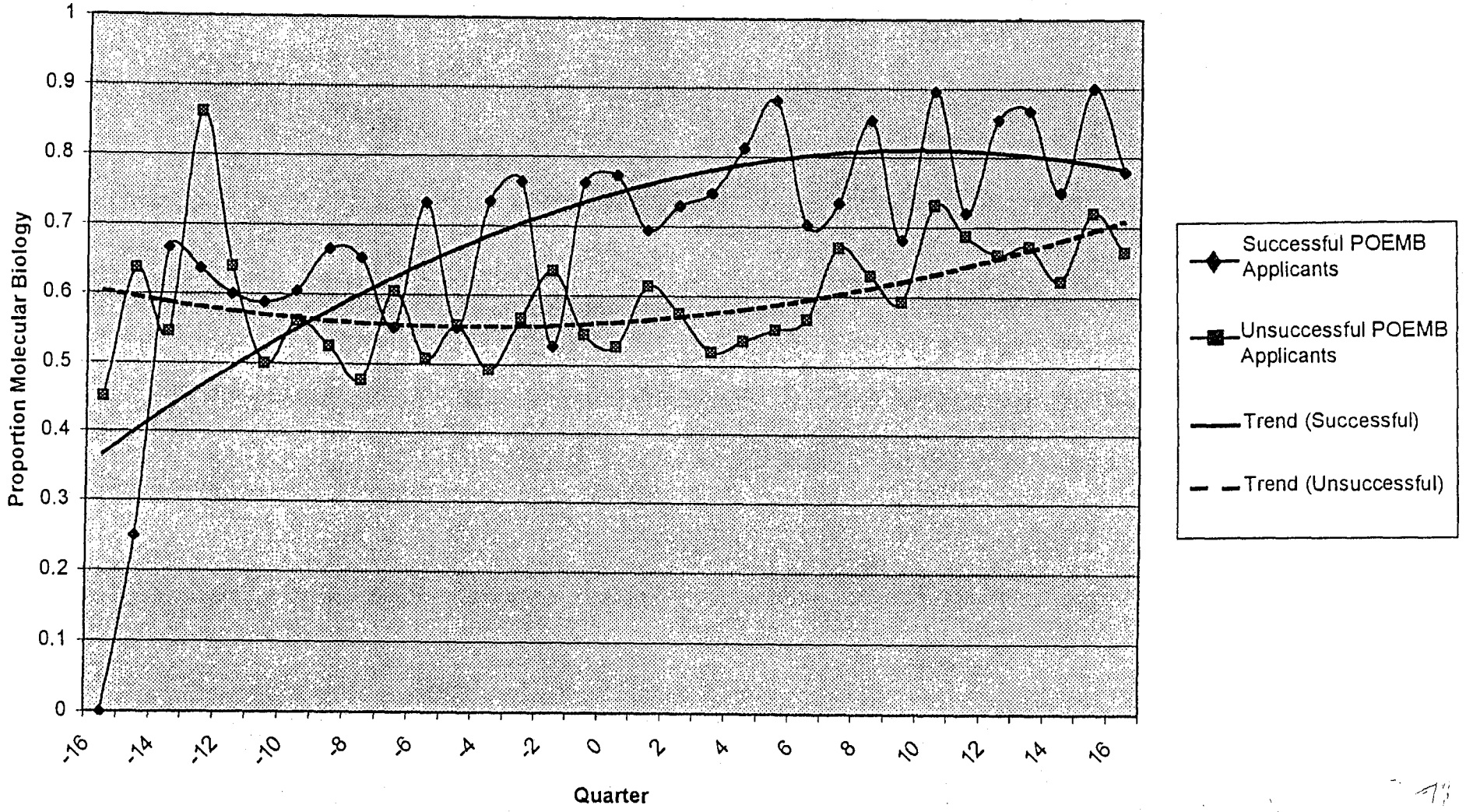
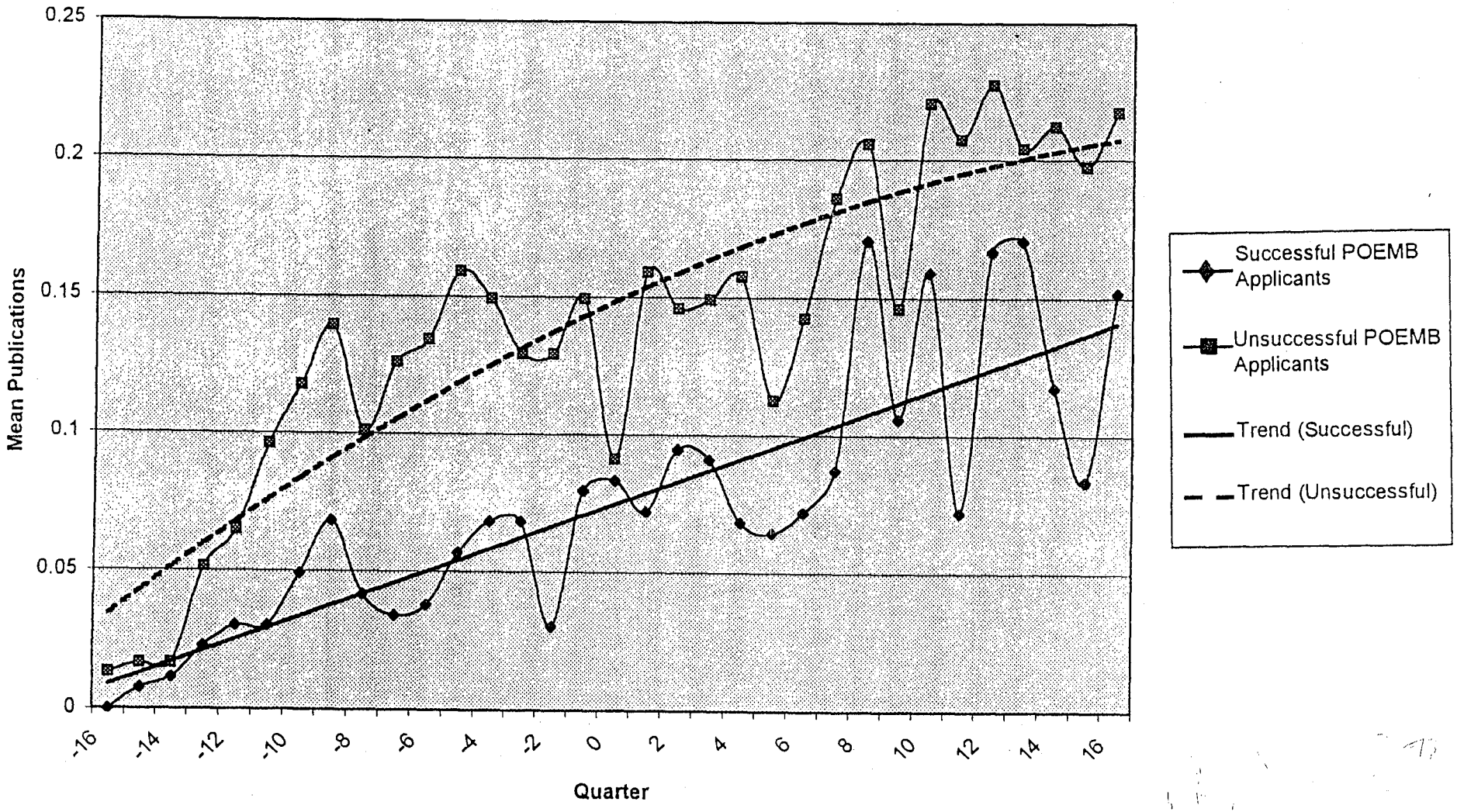


Figure 5.18 Mean Number of Journal Publications With Molecular Biology-Only MeSH Terms Per Investigator, By Quarter



patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

Figure 5.19 shows the ratio of articles with only molecular biology MeSH terms to total articles over time for each group. As shown, the groups are roughly equivalent through the pre-award period, while successful POEMB applicants are consistently higher through the post-award period. As with the analysis that included co-occurring terms, the before-after comparisons showed the successful POEMB applicants' average ratio to be significantly higher over the aggregate period ($p < 0.01$). As suggested by the figure, the pre-award difference (51% to 44%, not statistically significant) increased in the post-award period (66% to 48%, $p < 0.01$), consistent with an interaction effect. Among investigators with publications in both the pre- and post-award period, successful POEMB applicants gained 7 percentage points across the two periods, compared to 2 percentage for the unsuccessful applicants (difference not statistically significant due to low power). Stratification of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

Co-occurring Heart/Lung and Molecular Biology Terms. Figure 5.20 shows the average number of articles with co-occurring heart/lung and molecular biology-related MeSH terms over time for individuals in the successful and unsuccessful POEMB applicant groups, along with their associated trend lines. As shown, the data patterns are quite different than those observed when the two types of terms were examined separately. Unsuccessful POEMB applicants show a steady upward trend, while successful applicants show a similar trend initially which drops off for much of the post-award period before regaining much of the loss late in that period. The repeated measures analysis showed a main effect of time ($p < 0.01$), but no main effect of program, and an interaction effect of program by time in the linear term only ($p < 0.01$). Consistent with this, the before-after comparisons showed no significant difference in the aggregate or pre-award periods, but differences favoring the unsuccessful POEMB applicant group in the post-award periods and pre-post gain scores ($p < 0.01$). Examination by status showed that senior investigators from the successful POEMB applicant group had higher means for all periods. These differences were not statistically significant, however. A very similar pattern is apparent for the trainees in both POEMB applicant groups. (See the footnote in Section 5.2.1 for an explanation of this apparent paradox.)

Figure 5.19 Mean Proportion of Journal Publications With Molecular Biology Only MeSH Terms Per Investigator, By Quarter

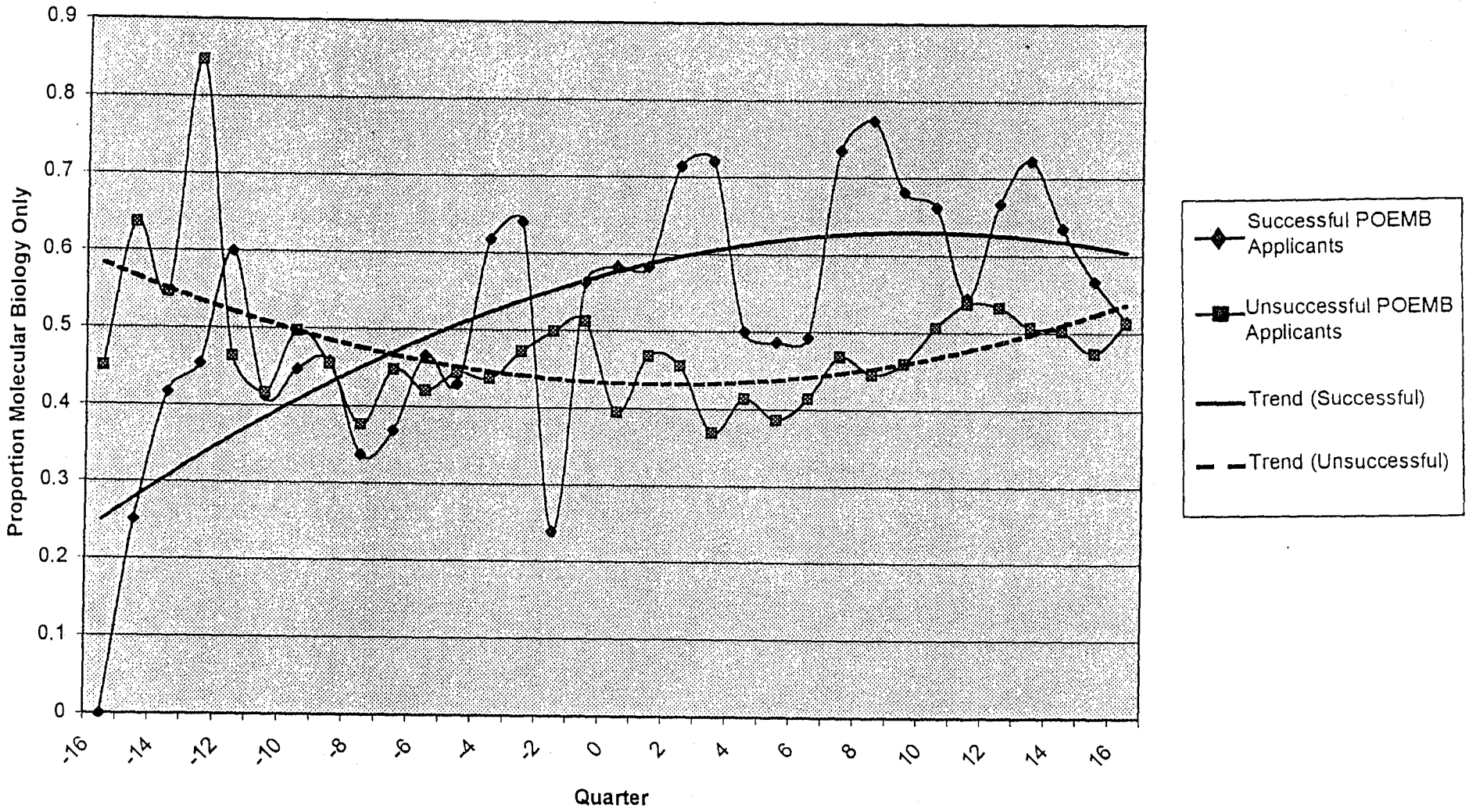


Figure 5.20 Mean Number of Journal Publications With Both Heart/Lung and Molecular Biology MeSH Terms Per Investigator, By Quarter

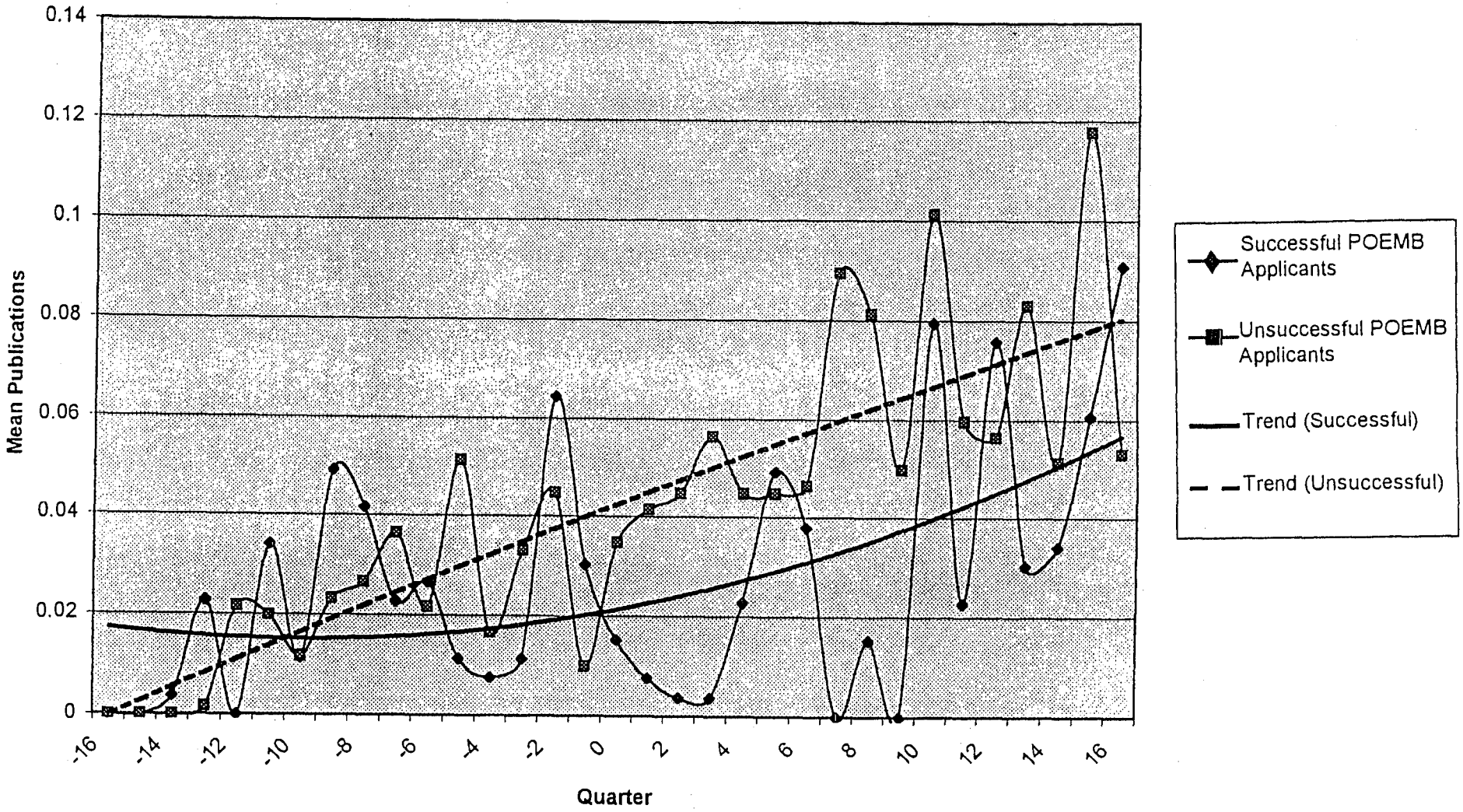


Figure 5.21 shows the ratio of articles with co-occurring heart/lung and molecular biology MeSH terms to total articles over time for each group. While high variability in the successful applicant series hinders interpretability somewhat, the pattern is clearly different than the prior series of proportions over time. There is no crossing of trend lines; rather, the successful POEMB applicants' trend line is above that of the unsuccessful applicants' trend line throughout the observation period, although the lines have practically converged by the end of the post-award period. As suggested by the figure, the pre-award difference (15% and 11% for successful POEMB applicants and unsuccessful applicants, respectively) decreased in the post-award period (18% and 16%); neither difference was statistically significant. Stratification of these same data by status revealed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

5.2.3.2 Analysis of "Switching"

Tables 5.5 and 5.6 show the complete transition matrices for successful and unsuccessful POEMB applicants, respectively. The transition matrices show the probability of all possible transitions from pre-POEMB to post-POEMB status. There are four possible pre- and post-states:[1]

- None - no **Medline** publications (and consequently no MeSH terms)
- HL only – heart, lung, or heart/lung MeSH terms only
- MB only - molecular/microbiology MeSH terms only
- HL/MB - HL and MB MeSH terms only

As with the previous switching analysis, each pre-POEMB state has a separate set of transition probabilities that sum to 100%. In Table 5.5, for example, among successful POEMB applicants who had no MeSH terms pre-award, 69% had no MeSH terms post-award, 22% had only MB MeSH terms, 7% had HL/MB MeSH terms, and the remaining 2% had only HL MeSH terms.

The darkly shaded cells represent the four transition states that could be categorized as "switching," so designated because each represents a transition from a non-HL to an HL state or a non-

---

[1] The reason there are only 4 states in the MeSH term analysis, compared to 8 in the **Medline** analysis, is the absence of an "other" category. The "other" category was dropped from the MeSH term analysis after discovering that every **Medline**-cited article by all 867 investigators had at least one MeSH descriptor classified as "other." This meant that the "other" classification was a constant rather than a variable, and would therefore add nothing meaningful to the analysis of switching.

Figure 5.21 Mean Proportion of Journal Publications With Both Heart/Lung and Molecular Biology MeSH Terms Per Investigator, By Quarter

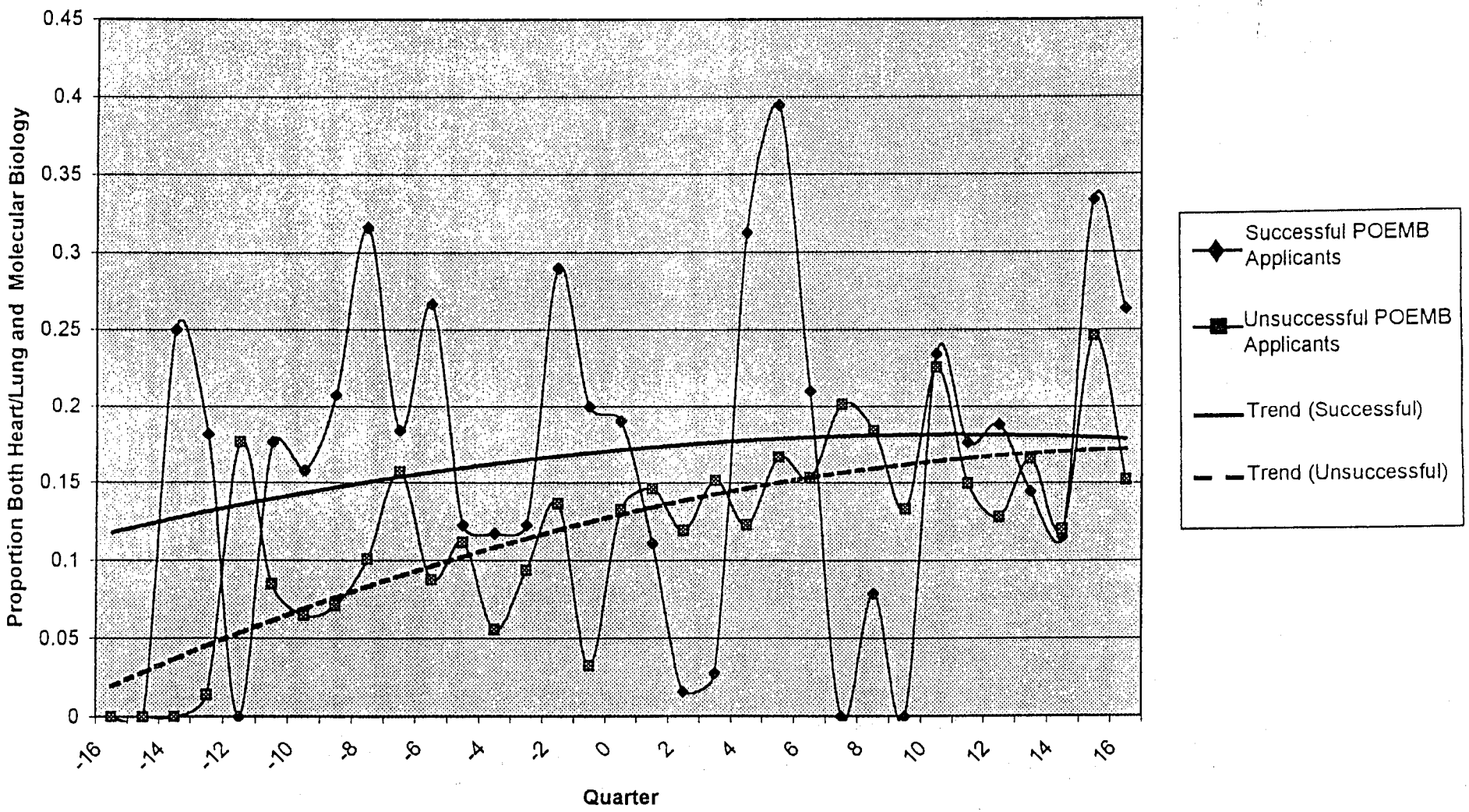


Table 5.5 Pre-POEMB Award to Post-POEMB Award MeSH Term Category Transition Matrix for Successful Applicants

Pre-POEMB Application	Post-POEMB Application				
Frequency Row Percent	HL	MB	HL/MB	none	Total Investigators
HL	1 17%	1 17%	1 17%	3 50%	6
MB	0 0%	17 61%	6 21%	5 18%	28
HL/MB	2 9%	4 17%	15 65%	2 9%	23
none	5 2%	45 22%	14 7%	143 69%	207
Total Percent	8	67	36	153	264

Table 5.6 Pre-POEMB Award to Post-POEMB Award MeSH Term Category Transition Matrix for Unsuccessful Applicants

Pre-POEMB Application	Post-POEMB Application				
Frequency Row Percent	HL	MB	HL/MB	none	Total Invesigators
HL	19 39%	1 2%	20 41%	9 18%	49
MB	3 2%	74 56%	42 32%	13 110%	132
HL/MB	7 6.%	17 16%	77 71%	7 6%	108
none	18 6%	73 23%	54 17%	169 54%	314
Total Percent	47	165	193	198	603

MB to an MB state. As with journal title categories, we characterized switching under two definitions: a more restrictive and a less restrictive. The more restrictive or "pure" switch conditions are those in which the investigator transitions from an MB to an HL state or vice versa. These are:

- HL only to MB only
- HL only to HL/MB
- MB only to HL only
- MB only to HL/MB

As shown in Tables 5.5 and 5.6, a relatively small number of investigators fell into any of the four "pure" switch transitions, and most were in the unsuccessful POEMB applicant group. Specifically, 2 were in the HL to MB group (1 unsuccessful applicant, 1 successful applicant), 21 were in the HL to HL/MB group (20 unsuccessful applicants, 1 successful applicant), 3 were in the MB to HL group (all unsuccessful applicants), and 48 were in the MB to HL/MB group (42 unsuccessful applicants, 6 successful applicants). In percent terms, the MB to HL/MB difference between groups was three-fold (7% to 2%), and statistically significant ($p < 0.01$).

The lightly shaded cells in the tables represent the additional less restrictive switch conditions. These were defined as follows:

- **No HL to some HL.** This includes none to HL only, none to HL/MB, MB only to HL only, MB only to HL/MB.
- **No MB to some MB.** This includes none to MB only, none to HL/MB, HL only to MB only, HL only to HL/MB.
- **No HL/MB to some HL/MB.** This includes none to HL/MB, HL only to HL/MB, MB only to HL/MB.

As in the journal title descriptor analysis, they are less restrictive in that investigators with and without pre-POEMB publication activity are grouped together. Of the 867 investigators, 142 met the No HL to some HL condition. Unsuccessful POEMB applicants were about twice as likely to fall into this subset as were successful applicants (19% to 9%), and this difference was statistically significant ($p < 0.01$). Among the subset of investigators who began in the No HL condition (and therefore had the opportunity to switch), the relative frequencies of the two groups showed a similar pattern (26% and 11% for unsuccessful and successful POEMB applicants, respectively).

More investigators (209) met the No MB to Some MB condition. Successful and unsuccessful POEMB applicants were about equally likely to fall into this group (23% and 24%, respectively, not statistically significant). Among the subset of investigators who began in the No MB condition, however, there was a pronounced difference favoring unsuccessful POEMB applicants (41% to 29%, $p < 0.01$).

As expected, a smaller number (137) met the No HL/MB to Some HL/MB condition. Unsuccessful POEMB applicants were substantially more likely to fall into this group (19% to 8%, $p < 0.01$). Among the subset of investigators who began in the No HL/MB condition, the pattern was similar (23% and 9% for the unsuccessful and successful POEMB applicants, respectively).

In sum, there were several switch patterns in which unsuccessful POEMB applicants transitioned at higher rates than successful applicants, others in which there was no significant difference, and none in which successful POEMB applicants transitioned at higher rates than unsuccessful applicants. Thus, these data offer no evidence that receiving a POEMB grant increased the probability of investigator switches of the sort hypothesized, at least as measured by the type of MeSH terms generated by the articles in which they published.

5.2.4 Journal Title Descriptor-MeSH Term Combinations

An additional approach to determining the effect of POEMB on the direction/focus of the research of individuals in both POEMB applicant groups is to characterize the journal articles using both the journal title descriptors and the MeSH terms. To this end, we examined the **Medline** data in terms of the percentage of investigators that published articles in:

- HL journals with MB MeSH terms
- MB journals with HL MeSH terms
- Cross-cutting journals with HL/MB MeSH terms
- "Other" journals with HL/MB MeSH terms

Regarding HL journal articles with MB MeSH terms, unsuccessful POEMB applicants were significantly more likely to achieve this in the aggregate, pre-, and post-award periods (all statistically significant at $p < 0.01$). The pre-award percentages were 5 percent and 1 percent for the unsuccessful POEMB applicants and successful applicants, respectively, and this difference increased to 9 percent and 1 percent during the post-award period. In addition, the likelihood of transitioning from no HL journal articles with MB MeSH terms in the pre-award period to one or more HL journal articles with MB MeSH

terms in the post-award period was 6 percent for the unsuccessful POEMB applicants and 1 percent for the successful applicants (also significant at $p < 0.01$). Finally, among the subset of investigators who began in the No HL Journal Articles with MB MeSH Terms condition (and therefore had the opportunity to transition), the pattern was identical (6 percent and 1 percent for the unsuccessful and successful POEMB applicants, respectively).

Regarding MB journal articles with HL MeSH terms, unsuccessful POEMB applicants were significantly more likely to achieve this in the aggregate and post-award periods (both statistically significant at $p < 0.01$), but not the pre-award. That is, on this indicator there is evidence that the unsuccessful POEMB applicants gained relative to the successful applicants across the pre- to post-interval. The pre-award percentages were 9 percent to 7 percent for the unsuccessful and successful applicants, respectively, and this difference increased to 16 percent to 9 percent during the post-award period. In addition, the likelihood of transitioning from no MB journal articles with HL MeSH terms in the pre-award period to one or more MB journal articles with HL MeSH terms in the post-award period was 12 percent for the unsuccessful POEMB applicants and 5 percent for the successful applicants (also significant at $p < 0.01$). Finally, among the subset of investigators who began in the No MB Journal Articles with HL MeSH Terms condition (and therefore had the opportunity to transition), the pattern was similar (13 percent and 5 percent for the unsuccessful and successful POEMB applicants, respectively).

Regarding cross-cutting journal articles with HL/MB MeSH terms, unsuccessful POEMB applicants were significantly more likely to achieve this in the aggregate ($p < 0.05$) and post-award ($p < 0.01$) periods, but not the pre-award. As with MB journal articles with HL MeSH terms, on this indicator there is evidence that the unsuccessful POEMB applicants gained relative to the successful applicants across the pre- to post-interval. The pre-award percentages were 4 percent and 5 percent for the unsuccessful and successful POEMB applicants, respectively, compared to 9 percent and 4 percent during the post-award period. In addition, the likelihood of transitioning from no cross-cutting journal articles with HL/MB MeSH terms in the pre-award period to one or more cross-cutting journal articles with HL/MB MeSH terms in the post-award period was 8 percent for the unsuccessful POEMB applicants and 2 percent for the successful applicants (also significant at $p < 0.01$). Finally, among the subset of investigators who began in the No Cross-Cutting Journal Articles with HL/MB MeSH Terms condition, the pattern was identical (8 percent and 2 percent for the unsuccessful and successful applicants, respectively).

Finally, regarding "other" journal articles with HL/MB MeSH terms, unsuccessful POEMB applicants once again were significantly more likely to achieve this in the aggregate ($p < 0.01$) and post-award ($p < 0.01$) periods, but not the pre-award. The pre-award percentages were 5 percent and 3 percent for the unsuccessful and successful applicants, respectively, compared to 15 percent and 5 percent during the post-award period. In addition, the likelihood of transitioning from no "other" journal articles with HL/MB MeSH terms in the pre-award period to one or more "other" journal articles with HL/MB MeSH terms in the post-award period was 13 percent for the unsuccessful POEMB applicants and 4 percent for the successful applicants (also significant at $p < 0.01$). Finally, among the subset of investigators who began in the No "Other" Journal Articles with HL/MB MeSH Terms condition, the pattern was identical (13 percent and 4 percent for the unsuccessful and successful POEMB applicants, respectively).

In sum, on all four indicators, unsuccessful POEMB applicants registered higher post-award percentages than did successful applicants, and in each case the difference between groups had increased relative to the pre-award period. From the standpoint of transitions, in all four cases unsuccessful POEMB applicants transitioned at significantly higher rates than did successful applicants. Thus, these data offer no evidence that receiving a POEMB grant increased the probability of publishing articles with the journal type-MeSH term combinations hypothesized.

5.3 Missionary Impact of POEMB on Those Investigators and Trainees Who Left the POEMB

This section examines the missionary impact of those who left the POEMB midway through the grant period. The section is subdivided into (1) methods and (2) findings.

5.3.1 Methods

In order to determine whether the investigators and trainees who left the POEMB continued to pursue applications of molecular biology in the fields of cardiovascular and pulmonary disease, their subsequent journal article publication activity was examined. Data utilized in the previous sections of this chapter to assess publication activity were used in this examination (i.e., **Medline** baseline file). The variables _1 through _7 from the **PMBPERS** file were used to identify those investigators and trainees who left the program and determine the timing of their departure. These seven variables represent

possible data sources for information contained in each record of **PMBPERS**: _1 (=the original application), _2 (=the first progress report), _3 (= the second progress report), etc. The fields that correspond to the document(s) in which the individual is cited are tagged with an "X" as a means of tracking POEMB investigators and trainees throughout the POEMB period. University of California at San Fransisco (UCSF) began their POEMB in 1989. The last year that we received information on these individuals was 1994. Thus, _6 was the last year that the UCSF individuals were tracked.

It should be noted that the category "other" used in the journal title designation includes those journals originally classified as "cross-cutting." It should also be noted that every article had a MeSH term distribution containing at least one "other"-designated term.

If the POEMB had an impact on the research careers (in terms of direction and focus) of those who departed the program before the end of the first funding period, we would expect to observe these individuals (1) publishing in journals classified as H/MB or L/MB and/or (2) publishing articles that are designated as H/MB/ "other", H/L/MB/ "other", and/or L/MB/ "other".

5.3.2 Findings

Table 5.7 presents information on the publication activity of those researchers who left POEMB before the end of the program.[1] A total of 126 researchers left their respective POEMBs before the end of the program, most of whom were trainees (SCODE 4). Most (51 percent) of the articles published by these individuals after their departure from the POEMB were classified by journal title as molecular biology-related. The next largest proportion (42 percent) of these articles were classified by journal title as "other"-related. In terms of MeSH term descriptor categories, 40 percent of the journal articles published after these individuals' departures were classified as heart- and/or lung- and molecular biology-related. This suggests that at least some of the individuals who left the POEMB before the end of the program continued to pursue applications of molecular biology in cardiovascular and/or pulmonary disease research.

---

1 Initially, publication output was stratified by institution. Examination of the publication activity of these individuals by status showed that individuals associated with one institution in particular were responsible for 88 percent of the H and/or L and MB articles as measured by MeSH terms. In the interest of preserving the confidentiality of the POEMB investigators, we present the publication data in aggregate.

Table 5.7 Post-Departure Publication Activity of Researchers Who Left POEMB Before the End of the Program

Number of Key Personnel and Trainees Who Left POEMB	Number of Journal Articles Published after Departure	Number of Articles with Journal Title Descriptors	Number of Articles Within Each of the Mesh Term Descriptor Categories*	
Status 1 =1	Status 1 = 0	H only = 5 (2.3%)	H/MB/ "other" = 17	(8%)
Status 2 = 4	Status 2 = 55	MB = 110 (51.2%)	H/L/MB/ "other" = 2	(1%)
Status 3 = 32	Status 3 = 72	Lung/MB = 9 (4.2%)	L/MB/ "other" = 66	(31%)
Status 4 = 89	Status 4 = 88	"Other" = 91 (42.3%)	H/ "other"=16	(7%)
Total = 126	Total = 215	Total = 215	MB/ "other" = 90	(42%)
			L/ "other" = 9	(4%)
			H/L/ "other" = 2	(1%)
			"Other" = 13	(6%)

* These values represent the totals across all statuses.

5.4 Summary of Findings

An examination of the data related to the total journal article publication activity of the two applicant groups shows that all investigators, regardless of status or applicant group, increase in publication activity from the pre- to the post-POEMB period. The most apparent pre- to post-POEMB patterns in publication activity are the almost four-fold increases in total number of articles published by the trainees in both the successful and unsuccessful applicant groups.

- On average, unsuccessful POEMB applicants published more than twice as many journal articles tracked by **Medline** as did successful POEMB applicants, a difference that was statistically significant ($p < 0.01$) for all analytical time periods. However, this difference in the pooled data was driven by the far higher percentage of trainees in the successful applicant group. In addition, the analysis of gain scores showed these differences increased over the pre-post period ($p < 0.01$).

- Examination of the data by status showed that for the total period, POEMB senior investigators had a higher mean number of publications per investigator than did the senior

investigators in the unsuccessful applicant group, a difference not statistically significant, due to low power.

- Both groups of senior investigators showed about the same amount of increase from pre- to post-, with the successful POEMB applicants starting higher and ending higher.
- A comparison between the trainees in the successful and unsuccessful applicant groups showed that these two groups were similar in terms of their mean number of publications for all periods.

5.4.1 Journal Title Descriptors

We looked at journal article publication activity using journal title descriptors to characterize overall disciplinary focus for (1) longitudinal analysis and (2) "switching" analysis.

5.4.1.1 Longitudinal Analysis

Certain general patterns emerged from the longitudinal analysis of the journal title descriptors. First, the unsuccessful applicants posted higher rates compared to the successful applicants in terms of the average number of publications over time per investigator in all five types of descriptor-designated journals. However, before-after comparisons of these same data showed more variability. Second, there was more variability in the differences between successful and unsuccessful applicants in terms of the ratios of specific descriptor-designated journal types to total publications for those investigators registering one or more publications in **Medline**. Third, in all but one case the stratification of trends by status (senior investigators versus trainees) showed patterns consistent with the results of the analyses in which all statuses were pooled within the applicant groups. Lastly, in only one case (molecular/microbiology-designated journals) did the evidence suggest that receiving a POEMB grant influenced the proportion of publications in a particular type of descriptor-designated journal.

Cardiovascular journals

- In terms of the average number of publications in cardiovascular journals per investigator, the difference between applicant groups was substantial and growing, with the publications by unsuccessful applicants exceeding those of the successful applicants, and notably larger than the difference between groups in total publication activity.
- In terms of the ratio of cardiovascular to total publications for those investigators registering one or more publications in **Medline**, the unsuccessful applicants' average ratio is consistently higher than that of the successful POEMB applicants, a difference that was significant ($p < 0.01$) for all periods.
- Thus, there was no evidence from these data that receiving a POEMB grant influenced the numbers or the proportion of publications in cardiovascular journals.

Respiratory journals

- The trends indicate increasing numbers of publications in respiratory journals for unsuccessful applicants relative to successful applicants throughout the pre-award period and beginning to decrease late in the post-award period.
- Before-after comparisons showed six-fold or greater differences between groups for the aggregate, pre-, and post-periods (all significant at $p < 0.01$), but neither group showed any pre-post gains.
- As with cardiovascular journals, examination of the ratio of respiratory to total showed that in most intervals the unsuccessful POEMB applicants' ratio is higher.
- Before-after comparisons showed differences favoring unsuccessful applicants for the aggregate, pre-award, and post-award periods, though none were statistically significant.
- Among investigators who published in both periods, the respiratory-to-total ratio increased by a greater amount among successful applicants, but the difference was small and nonsignificant.
- Stratification of these data by status revealed a pattern consistent with the results of the analysis in which all statuses were pooled within the applicant groups.
- Overall, there was no evidence from these data that receiving a POEMB grant influenced the numbers or the proportion of publications in respiratory journals.

Molecular/microbiology journals

- The trends indicate both groups increasing their publication in molecular/microbiology journals steadily over time, with the unsuccessful applicants producing more articles than the successful applicants throughout most of the pre- and post-award periods.

- The repeated measures analysis showed main effects of program, time, and time-squared (all at $p < 0.01$), no interaction effect of program by time in the linear or term, but an interaction in the quadratic term ($p < 0.05$) consistent with the unsuccessful applicant group's curvilinearity.

- Before-after comparisons showed differences between groups for the aggregate, pre-, and post-periods (all significant at $p < 0.01$), but no differences in the pre-post gains, which were substantial in both groups.

- Examination of the data by status showed the POEMB senior investigators with higher mean numbers of publications in molecular/microbiology-designated journals for all periods, and increasing to a greater degree in the post-POEMB period. The total and post-period differences were statistically significant. A very similar pattern is apparent for the trainees.

- In terms of the ratio of articles in molecular/microbiology journals to total journal output, the pattern differs from those for the cardiovascular and respiratory journals: the successful applicants began lower than their counterparts, draw even in the latter part of the pre-award period, and pass them in the post-award period.

- Before-after comparisons of the ratios show the successful applicants with a higher ratio over the aggregate period ($p < 0.01$). The post-award difference, favoring the successful applicants, was substantial and statistically significant ($p < 0.01$). This suggests a group by time interaction and is consistent with an explanation that POEMB played a role in the shift in publication emphasis to molecular/microbiology journals over time.

- Examination of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

- Overall, there is evidence that receiving a POEMB grant influenced the proportion of publications in molecular/microbiology journals.

Cross-cutting journals

- The trends show unsuccessful applicants publishing more articles in cross-cutting journals throughout the observation period, with little change in the size of the difference over time.

- The repeated measures analysis showed main effects of program ($p < 0.01$) and time-squared ($p < 0.05$), but no interaction effect of program by time in the linear or quadratic terms.

- Before-after comparisons showed differences between groups for the aggregate ($p < 0.01$), pre ($p < 0.05$), and post-periods ($p < 0.01$). The comparison of pre-post gain scores showed that unsuccessful POEMB applicants increased their numbers of publications in cross-cutting journals relative to successful applicants, but the difference was nonsignificant.

- In terms of the ratio of cross-cutting to total publications over time for each group, the trends are somewhat difficult to interpret due to high variability over time in the successful POEMB applicant group.

- Before-after comparisons of the ratio showed successful applicants with a significantly higher mean ratio during the pre-award period ($p < 0.05$), which disappeared and changed direction in the post-award period (not statistically significant).

- Examination of these same data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

- Thus, receiving a POEMB grant was associated with a decrease in the mean proportion of articles appearing in cross-cutting journals.

"Other" journals

- The trends show that, after beginning at approximately the same level, the unsuccessful applicants' series increases steadily while the successful applicants' series remains flat until late in the post-award series.

- The repeated measures analysis showed main effects of program, time, and time-squared, and interaction effects of program by time in both the linear and quadratic terms (all significant at $p < 0.01$).

- Before-after comparisons showed two-fold or greater differences between applicant groups for the aggregate, pre-, and post-periods, as well as still larger relative differences (six-fold) in gains over time (all significant at $p < 0.01$).

- In terms of the "other" to total ratio over time for each group, the overall trends indicate little change over time and little difference between groups.

- Before-after comparisons of the ratios showed no significant differences over the aggregate, pre-award, or post-award periods, and no differences in gain scores.

- Examination of the data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

- Thus, there is no evidence from these data that receiving a POEMB grant influenced the numbers or the proportion of publications in "other" journals.

5.4.1.2 Analysis of "Switching"

There were several switch patterns in which unsuccessful POEMB applicants transitioned at higher rates than did successful applicants, others in which there was no significant difference, and none in which successful POEMB applicants transitioned at higher rates than did unsuccessful applicants. Thus, these data offer no evidence that receiving a POEMB grant increased the probability of investigator switches of the sort hypothesized, at least as measured by the disciplinary focus of the journals in which they published.

5.4.2 Article Descriptors—MeSH Terms

We looked at journal article publication activity using MeSH term categories to characterize overall focus of the articles for (1) longitudinal analysis and (2) "switching" analysis.

5.4.2.1 Longitudinal Analysis

Certain general patterns emerged from the longitudinal analysis of the MeSH term article descriptors. First, the unsuccessful applicants tended to post higher rates compared to the successful applicants in terms of the average number of publications over time per investigator for all three types of MeSH term-designated articles. The one exception to this pattern is found in the trend over time for the average number of journal publications with *only* molecular biology MeSH terms. Second, there was considerable variability in the differences between successful and unsuccessful applicants in terms of the ratios of specific MeSH term-designated journal publications to total publications. Third, in all but one case (co-occurring heart/lung and molecular biology terms) the stratification of trends by status (senior investigators versus trainees) showed patterns consistent with the results of the analyses in which all statuses were pooled within the applicant groups. Lastly, in only one case (molecular biology-designated journal publications) did the evidence suggest that receiving a POEMB grant influenced the proportion of publications in a particular type of MeSH term-designated journal.

Findings are reported first for journal publications with MeSH terms that co-occur with other MeSH terms, as well as for journal publications with only MeSH terms relevant to the analyses carried out here (i.e., no co-occurrence with other MeSH terms).

Heart/lung terms

- In terms of the average number of journal publications with *heart/lung MeSH terms* per investigator, articles with heart/lung MeSH terms begin at similar levels in both groups, then diverge about midway through the pre-award observation period (with unsuccessful applicants increasing faster), and continuing to diverge throughout much of the post-award period.

- The repeated measures analysis showed main effects of program, time (both at $p < 0.01$), and time-squared ($p < 0.05$), and an interaction effect of program by time in the linear term ($p < 0.01$), but no interaction in the quadratic term.

- Before-after comparisons showed differences between groups for the aggregate, pre, and post periods, as well as large differences in gains over time (all significant at $p < 0.01$).

- In terms of the ratio of articles generating heart/lung-designated MeSH terms to total publications, the crossing of the trend lines suggests a group by time interaction, in which the unsuccessful applicants' ratio is below that of the successful applicants' through much of the pre-award period, but higher thereafter.

- Before-after comparisons of the ratio revealed significantly higher proportions among unsuccessful applicants over the aggregate period ($p < 0.05$). There was no difference over the pre-award period, but a difference over the post-award period ($p < 0.05$), consistent with an interaction effect.

- In terms of the average number of journal publications with *heart/lung MeSH terms only* the difference between applicant groups is even more pronounced, with successful POEMB applicants showing virtually no upward trend.

- The repeated measures analysis showed main effects of program, time, and time-squared (all at $p < 0.01$), and interaction effects of program by time in the linear ($p < 0.01$) and quadratic ($p < 0.05$) terms.

- Before-after comparisons showed differences between groups for the aggregate, pre, and post periods, with the size of the difference more than doubling over the pre-post periods (all significant at $p < 0.01$).

- In terms of the ratio of articles with *only* heart/lung MeSH terms to total articles, the trend lines suggest an even more pronounced interaction effect than when co-occurring terms were included.

- Before-after comparisons revealed significantly higher proportions among unsuccessful POEMB applicants for the aggregate, pre-, and post-award periods, as well as a significant difference in gain scores ($p < 0.01$), consistent with an interaction. There was no difference over the pre-award period, but a difference over the post-award period ($p < 0.01$), consistent with an interaction effect.

- Overall, there was no evidence from the data that receiving a POEMB grant influenced the numbers or the proportion of journal publications with heart/lung MeSH terms, neither those as the only term nor co-occurring with other terms.

Molecular biology terms

- In terms of the average number of journal publications with *molecular biology MeSH terms* over time per investigator, both groups increased over time on this dimension, with the unsuccessful applicants increasing at a somewhat faster pace.

- The repeated measures analysis showed main effects of program and time (both at $p < 0.01$) and an interaction effect of program by time in the linear ($p < 0.01$) and quadratic ($p < 0.05$) terms.

- Before-after comparisons showed differences between groups for the aggregate, pre-, and post-periods (all significant at $p < 0.01$), with a smaller yet still significant ($p < 0.05$) difference in gain scores.

- Stratification of these same data by status revealed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

- In terms of the ratio of articles with molecular biology MeSH terms to total articles over time for each applicant group, the ratio is higher among successful applicants for most of the pre-award period and all of the post-award period. Note also that the proportions in both groups are 0.5 or higher in almost every period, indicating that the majority of articles published by these investigators during the observation period generated molecular biology MeSH terms.

- Before-after comparisons of the ratio of articles with molecular biology MeSH terms to total articles revealed that successful applicants' average proportion was significantly higher over the aggregate period ($p < 0.01$). The pre-award difference ($p < 0.05$) increased in the post-award period ($p < 0.01$).

- In terms of the average number of journal publications with *only molecular biology MeSH terms* per investigator, successful applicants increased at a slightly more comparable rate than when articles with co-occurring heart/lung terms were lumped in.

- The repeated measures analysis showed main effects of program, time (both at $p < 0.01$), and time-squared ($p < 0.05$), and an interaction effect of program by time in the linear term ($p < 0.05$), but no interaction in the quadratic term.

- Before-after comparisons still showed differences between groups for the aggregate, pre, and post periods (all significant at $p < 0.01$), but no differences in the pre-post gain scores.

- In terms of the ratio of articles with only molecular biology MeSH terms to total articles over time for each group, the groups are roughly equivalent through the pre-award period, while successful applicants are consistently higher through the post-award period.

- Before-after comparisons of the ratio showed, as with the analysis that included co-occurring terms, the successful applicants' average ratio to be significantly higher over the aggregate period ($p < 0.01$). The pre-award difference (not statistically significant) increased in the post-award period ($p < 0.01$), consistent with an interaction effect. Among investigators with publications in both the pre- and post-award periods, successful applicants gained at a greater rate across the two periods, compared to the unsuccessful applicants (difference is not statistically significant due to low power).

- Overall, there was evidence from the data that receiving a POEMB grant influenced the proportion of journal publications with molecular biology MeSH terms (co-occuring), and those with only molecular biology MeSH terms.

Co-occurring heart/lung and molecular biology terms

- In terms of the average number of articles with co-occurring heart/lung and molecular biology-related MeSH terms over time per individuals in both applicant groups, the data patterns are quite different than those observed when the two types of terms were examined separately. Unsuccessful POEMB applicants show a steady upward trend, while successful applicants show a similar trend initially but one that drops off for much of the post-award period before regaining much of the loss late in that period.

- The repeated measures analysis showed a main effect of time ($p < 0.01$), but no main effect of program, and an interaction effect of program by time in the linear term only ($p < 0.01$).

- Before-after comparisons showed no significant difference in the aggregate or pre-award periods, but differences favoring the unsuccessful POEMB applicant group in the post-award periods and pre-post gain scores ($p < 0.01$).

- Examination by status showed POEMB senior investigators higher for all periods (but not statistically significant). POEMB trainees were slightly higher for total period but lower in post-period (not statistically significant).

- In terms of the ratio of articles with co-occurring heart/lung and molecular biology MeSH terms to total articles over time for each group, high variability in the successful applicant series hinders interpretability somewhat but the pattern is clearly different than the prior MeSH term series of proportions over time. There is no crossing of trend lines; rather, the successful POEMB applicants' trend line is above that of the unsuccessful applicants' trend line throughout the observation period, although the lines have practically converged by the end of the post-award period.

- Before-after comparisons of the ratio showed that the successful applicants' greater pre-award ratio decreased in the post-award period (not statistically significant).

- Stratification of these same data by status revealed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

- Overall, the evidence does not strongly suggest, due to the narrowing differences between applicant groups in the post-award period, that receiving a POEMB grant influenced the number or proportion of journal publications with co-occuring heart/lung and molecular biology MeSH terms.

5.4.2.2 Analysis of "Switching"

There were several switch patterns in which unsuccessful POEMB applicants transitioned at higher rates than successful applicants, others in which there was no significant difference, and none in which successful POEMB applicants transitioned at higher rates than unsuccessful applicants. Thus, these data offer no evidence that receiving a POEMB grant increased the probability of investigator switches of the sort hypothesized, at least as measured by the type of MeSH terms generated by the articles in which they published.

5.4.3 Journal Title Descriptor—MeSH Term Combinations

Research journal articles were characterized using both the journal title descriptors and the MeSH terms to determining the effect of POEMB on the direction/focus of the research of individuals in both POEMB applicant groups. To this end, we examined the **Medline** data in terms of the percentage of

investigators that published articles in: HL journals with MB MeSH terms; MB journals with HL MeSH terms; cross-cutting journals with HL/MB MeSH terms; "other" journals with HL/MB MeSH terms.

On all four indicators, unsuccessful applicants registered higher post-award percentages than did successful applicants, and in each case the difference between groups had increased relative to the pre-award period. From the standpoint of transitions (switching), in all four cases unsuccessful POEMB applicants transitioned at significantly higher rates than did successful applicants. Thus, these data offer no evidence that receiving a POEMB grant increased the probability of publishing articles with the journal type-MeSH term combinations hypothesized.

5.4.4 Missionary Impact of POEMB

Based on the publication activity of those researchers who left their respective POEMB before the end of the program (T=126) we found that:

- Most (51 percent) of the articles published by these individuals after their departure from the POEMB were classified by journal title as molecular biology-related.
- The next largest proportion (42 percent) of these articles were classified by journal title as "other"-related.
- In terms of MeSH term descriptor categories, 40 percent of the journal articles published after these individuals' departures were classified as heart- and/or lung- and molecular biology-related.
- These findings suggests that at least some of the individuals who left the POEMB before the end of the program continued to pursue applications of molecular biology in cardiovascular and/or pulmonary disease research.

- These findings suggests that at least some of the individuals who left the POEMB before the end of the program continued to pursue applications of molecular biology in cardiovascular and/or pulmonary disease research.

5.5 Factors Limiting Analysis

The factors influencing the **Medline** portion of our analysis are similar to those influencing the grant application and award analyses. The **Medline** baseline file contained no articles published after 1993, which does not take us to the end of the first POEMB grant period. Because of the well-known lag time between, conduct of research and the publication of findings, it is likely we are not yet able t to see POEMB impact on publication patterns. Furthermore, POEMB researchers were encouraged to engage in more innovative research on complex experimental systems and models—research that requires additional time to develop and apply. This means that the earliest observable POEMB impact might in fact be a drop in publication activity among investigators from the successful applicant institutions. Future analysis with this data set using an extended **Medline** file might be better able to detect changes in publication activity that might be attributable to the effect of the POEMB. Use of the *Science Citation Index* in this analysis would help determine whether the POEMB publications are being widely cited by other researchers in the field.

progress reports or on the part of the Battelle staff who reviewed the documents. Overall, the POEMB institutions were not consistent in terms of their record keeping which was reflected in the progress reports. In addition, because the University of California at San Francisco (UCSF) began their POEMB in 1989, progress reports from this institution were only available for the first five years of the program. Thus, the tracking period for this institution is not fully comparable to those POEMB institutions funded in 1988. For these reasons, this variable was not used to characterize the investigators and trainees from the applicant groups in terms of the timing of their entry into the POEMB program.

In an approach consistent with the goal of assessing whether the POEMB affected the research direction/focus of individuals in the successful POEMB applicant group, this analysis focuses mainly on the investigator (or individual researcher) as the unit of analysis. The *dependent variables* considered in this portion of the analysis that came directly from the **CGAF/CRISP** baseline file include (1) total NIH/PHS grant applications and (2) total applications classified as heart/lung (HL) using ICD codes. Other dependent variables were constructed from information contained in the **CGAF/CRISP** baseline file. These include (3) the ratio of HL-designated applications to total applications, (4) having submitted one or more total applications versus having submitted no applications, and (5) having submitted one or more HL-designated applications versus having submitted no HL-designated applications. Quarters are considered the unit of time in all longitudinal analyses discussed in this chapter. The timing of the applications is involved in the hypothesis testing (i.e., the examination of the grant application data for group by time interactions).

The analyses focused on the following measures of the differences in the grant application activity of researchers in the two POEMB applicant groups:

- Aggregate number of total grant applications submitted in the pre-POEMB award, post-POEMB award, and total periods.
- Mean number of total and HL-designated grant applications submitted in the pre-POEMB award, post-POEMB award, and total periods.
- Repeated-measures analysis of the above-noted dependent variables.
- HL-designated (ICD) applications as a proportion of total applications in the pre-POEMB award, post-POEMB award, and total periods.
- Visual presentation of each dependent variable over time (17 quarters before to 16 quarters after POEMB application), with trend lines fit to the data.

- Transition probabilities from all possible pre-POEMB states to all possible post-POEMB states.
- Proportional differences in terms of "switches" (changes from one ICD classification state to another).

These differences were examined by pooling all status types (SCODEs 1, 2, 3, and 4), and then comparisons were made between the senior investigators (SCODEs 1 and 2) in the two POEMB applicant groups as well as between trainees and potential trainees (SCODEs 3 and 4 for successful applicants are compared to SCODE 3 for unsuccessful applicants). It was important to break out senior investigators and trainee-level researchers because the program was expected to affect senior investigators and trainees differently on criterion outcomes, and also because the percentage of senior investigators was far lower in the successful than in the unsuccessful applicant group. Specifically, 13 percent of successful applicants were senior investigators compared to 63 percent of unsuccessful applicants. The highly unbalanced distributions of senior investigators versus trainee-level researchers across the two applicant groups can lead to misleading results when the senior investigators and trainees are pooled. In the most extreme case — a phenomenon known as Simpson's paradox (Bickel et al., 1975) — the data can show a consistent result across both subgroups (e.g., both senior investigators and trainees submitting more applications on average in the successful applicant group) while the pooled data show the opposite result (all investigators submitting fewer applications on average in the successful applicant group).

In all analyses discussed in this chapter, status 3s (potential trainees) and 4s (trainees cited in POEMB progress reports) in the successful POEMB applicant group are combined as "trainees". We recognize that conceptually there are no individuals in the unsuccessful POEMB applicant group that are truly comparable to the status 4s in the successful applicant group. Yet it is the status 4s that are of particular interest to us because of the evidence (citation in a POEM progress report) that these researchers were actually supported for a period of time by the POEMB grant[2]. In order to investigate

---

[2] Status 3s appear in a POEMB application but not in a subsequent progress report, so evidence that they actually worked on the grant is more tenuous than for status 4s. We posit four possible relationships of a status 3 individual to the POEMB grant: (1) the status 3 individual may never have worked on the POEMB grant and been merely a "placeholder" in the application for a trainee-level individual to be recruited; (2) the status 3 individual may have worked briefly on the grant but left before the first progress report was compiled; (3) the status 3 individual may have worked on the grant but not been cited by the progress report compiler due to clerical oversight;

the potential bias introduced by including the status 4s in the "trainee" comparisons, we also compared the equivalency of the status 3s and 4s in the successful POEMB applicant group on all outcomes of interest. Cases in which the exclusion of the status 4s may have made a difference in the results and interpretations are cited in the text of this chapter.

Visual comparisons over the entire pre-post-period (longitudinal) are also an important focus of this analysis. These comparisons are provided in the form of plots over time and trend lines to fit plots.

In all cases where the data support it, statistical modeling was carried out on the full sample (senior investigators and trainees pooled) in the form of repeated measures analysis. All such analyses were conducted using PROC MIXED in SAS®, with POEMB as a fixed effect, investigator as a random effect nested within POEMB, the dependent variables (separately) as polynomial functions of time (17 quarters before to 16 quarters after POEMB application), and the covariance structure of repeated observations determined empirically (compound symmetry or autoregressive). The repeated measures analysis permits us to statistically test the hypothesis that receiving a POEMB award influenced the dependent variables in question. It is superior to a simple comparison of pre-post gain scores because: (1) it uses many more observations (quarterly data), thereby enabling the detection of trends that would not be apparent from pre-post gain scores in which those observations are aggregated (e.g., quadratic effects of time), and (2) it is more statistically powerful. The particular repeated measures model used, a mixed model in which the covariance structure of the quarterly observations was determined empirically, provides the most accurate probability levels for significance tests of model parameters (Littell et al., 1996).

The main (null) hypothesis tested in this section is that there are no (statistically significant) group by time interactions. Given the nature of the POEMB, we would expect that being awarded the POEMB would have a "shielding" effect on the senior investigators (SCODEs 1 and 2), in other words, these investigators would be freed from having to submit multiple grant applications to maintain a stream of funding for their research, which would presumably allow them to spend more of their time on research. This effect may manifest itself as a decrease in the total grant activity of the POEMB senior investigators during the post-POEMB application period. The research activity of POEMB trainees and

---

or (4) the status 3 individual may have worked on the grant and been cited in a progress report, but we overlooked the citation. The status 3 individual thus may or may not have been in a position to be directly influenced by the POEMB.

potential trainees (SCODEs 4 and 3, respectively), however, may have been stimulated by the program, leading them to pursue funding (submit grant applications) for more new research projects. We would also expect that being awarded the POEMB may have had an impact on the research direction/focus of the individuals in the successful POEMB applicant group. In other words, receiving the POEMB may have led to a shift or switch from a focus on noncardiovascular/pulmonary research (as measured by grant application activity and ICD) to cardiovascular/pulmonary research. It is also conceivable that a similar effect might be observed even among unsuccessful applicants if the collaborative institutional effort required to prepare a POEMB proposal has stimulated researchers' interest in such cross-disciplinary work.

There is an important limitation to the analyses reported on in this section. The range of the **CGAF** file (up to 1993) did not allow us to look at the entire period of the first POEMBs (1988 to 1995 for MIT and U Cincinnati; 1989 to 1996 for UCSF). This may have caused us to miss important changes in the research activity of the individuals included in this study — changes that may have taken place between 1993 and the end of the first funding period.

4.1.2 Findings

Below we discuss findings related to NIH/PHS grant activity in terms of (1) longitudinal analysis and (2) "switching" analysis.

4.1.2.1 Longitudinal Analysis

Using the **CGAF** baseline file, we first characterized NIH grant application activity in terms of the total number of grant applications submitted by successful (N=264) and unsuccessful POEMB applicants (N=603) during the period 17 quarters before and 16 quarters after POEMB application.[3] Table 4.1 presents the frequency of applications submitted by both POEMB applicant groups by status (SCODE) during the period 17 quarters before and 16 quarters after POEMB application. Investigators

[3] Additional information obtained from the **CGAF** baseline file allowed for the identification of four additional duplicates in the **PMBPERS** file. These four duplicates were removed subsequent to the analyses carried out in Chapter 3.

in the successful POEMB applicant group submitted a total of 126 grant applications during this time period. A total of 1,232 grant applications were submitted by unsuccessful POEMB applicants during this same period. The first apparent pre-post POEMB difference in grant application activity is the decrease in total number of applications submitted by senior investigators (SCODEs 1 and 2) in the successful POEMB applicant group (49 to 33). Note that a similar decrease among senior investigators was *not* observed in the unsuccessful applicant group. This pattern is consistent with the hypothesis that POEMB shields these investigators from having to apply for multiple grants to maintain funding streams. The second difference is the more than two-fold increase in total submissions (10 to 25) exhibited by the trainees (SCODE 4) from the successful POEMB applicant group. At first glance, this might appear to represent "maturation," i.e., the fact that junior investigators will become more active in submitting grant applications as their research careers develop. Had the effect been simply due to maturation, a similar increase would be observed among potential trainees (status 3) in the unsuccessful group; however, it was not. Thus, this difference is consistent with the hypothesis that being awarded the POEMB afforded trainees the opportunity and provided the stimulus to pursue other grants independently.

On average (mean number of total applications as opposed to aggregate total), unsuccessful POEMB applicants (SCODEs 1, 2, 3, or 4) submitted more than four times as many total applications as did successful POEMB applicants, and almost four times as many NHLBI applications. These relationships held over the total period, pre-POEMB period, and post-POEMB period, and each contrast was statistically significant ($p < 0.01$). Examination of these same data by status code (SCODE) eliminated the four-fold difference in grant submission between the successful and unsuccessful POEMB applicants for all three periods (total, pre and post). That is, the four-fold difference in the pooled data was largely driven by the far higher percentage of trainees in the successful applicant group (87 percent versus 37 percent), as described in the methods section. Although much reduced by stratification, the difference between the average number of total grants submitted by senior investigators (SCODEs 1 and 2) in the two applicant groups remained statistically significant ($p < 0.01$) for the total period, with successful POEMB applicants submitting fewer grant applications overall. For the pre-POEMB period, a comparison of successful and unsuccessful POEMB applicants in this same stratum showed no

Table 4.1 NIH/PHS Grant Applications Submitted by Investigators in the Successful and Unsuccessful POEMB Applicant Groups for the Period 17 Quarters Before and 16 Quarters After POEMB Application

Investigators from Successful POEMB Applicant Group	Applications Submitted 17 Quarters Pre-POEMB Row %, Column %	Applications Submitted 16 Quarters Post-POEMB Row %, Column %	Total
Status 1 and 2 (Status 1, N=4; Status 2, N=30)	49 60% 78%	33 40% 52%	82
Status 3 (N=32)	4 44% 6%	5 56% 8%	9
Status 4 (N=198)	10 29% 16%	25 71% 40%	35
Investigators from Unsuccessful POEMB Applicant Group			**Total**
Status 1 and 2 (Status 1, N=18; Status 2, N=365)	526 51% 86%	512 49% 82%	1038
Status 3 (N=220)	84 43% 14%	110 57% 18%	194
Total	673	685	1358

difference in the average number of grants submitted. A post-POEMB comparison revealed that the average number of grants submitted by successful POEMB senior investigators decreased relative to the senior investigators in the unsuccessful applicant group. This difference was statistically significant ($p<0.01$) and is consistent with the hypothesis that POEMB shields the senior investigators in the successful applicant group from having to apply for multiple grants to maintain funding streams.

Figures 4.1 and 4.2 plot these activities by quarter from 16 quarters prior to the POEMB application date to 16 quarters post, along with the least squares quadratic trend line through each time series. As shown, the trends show little or no change in level from pre to post. Consequently there is no indication of either absolute or relative change over time.

Figure 4.1 Mean Number of Applications, by Quarter (Total)

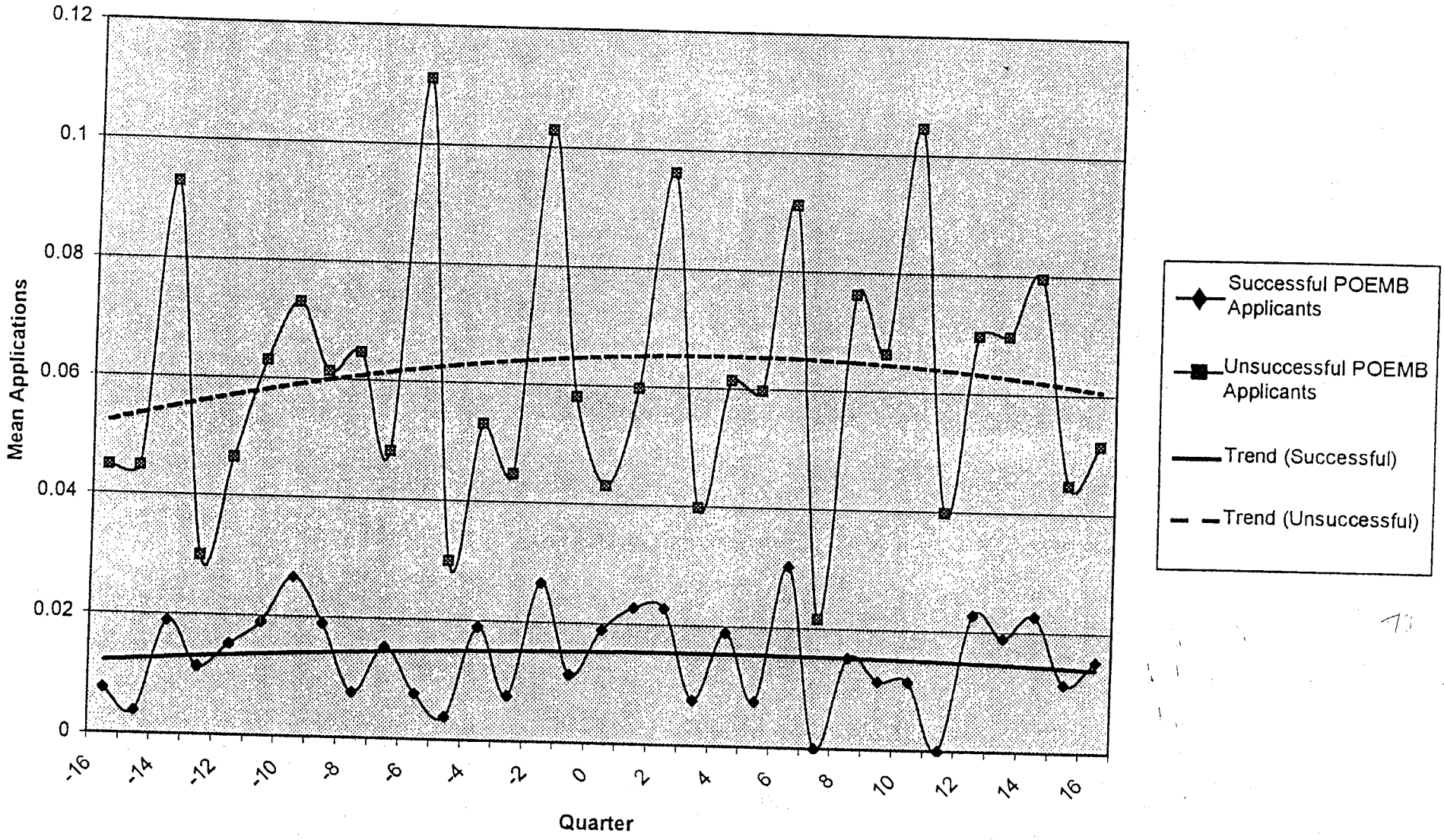
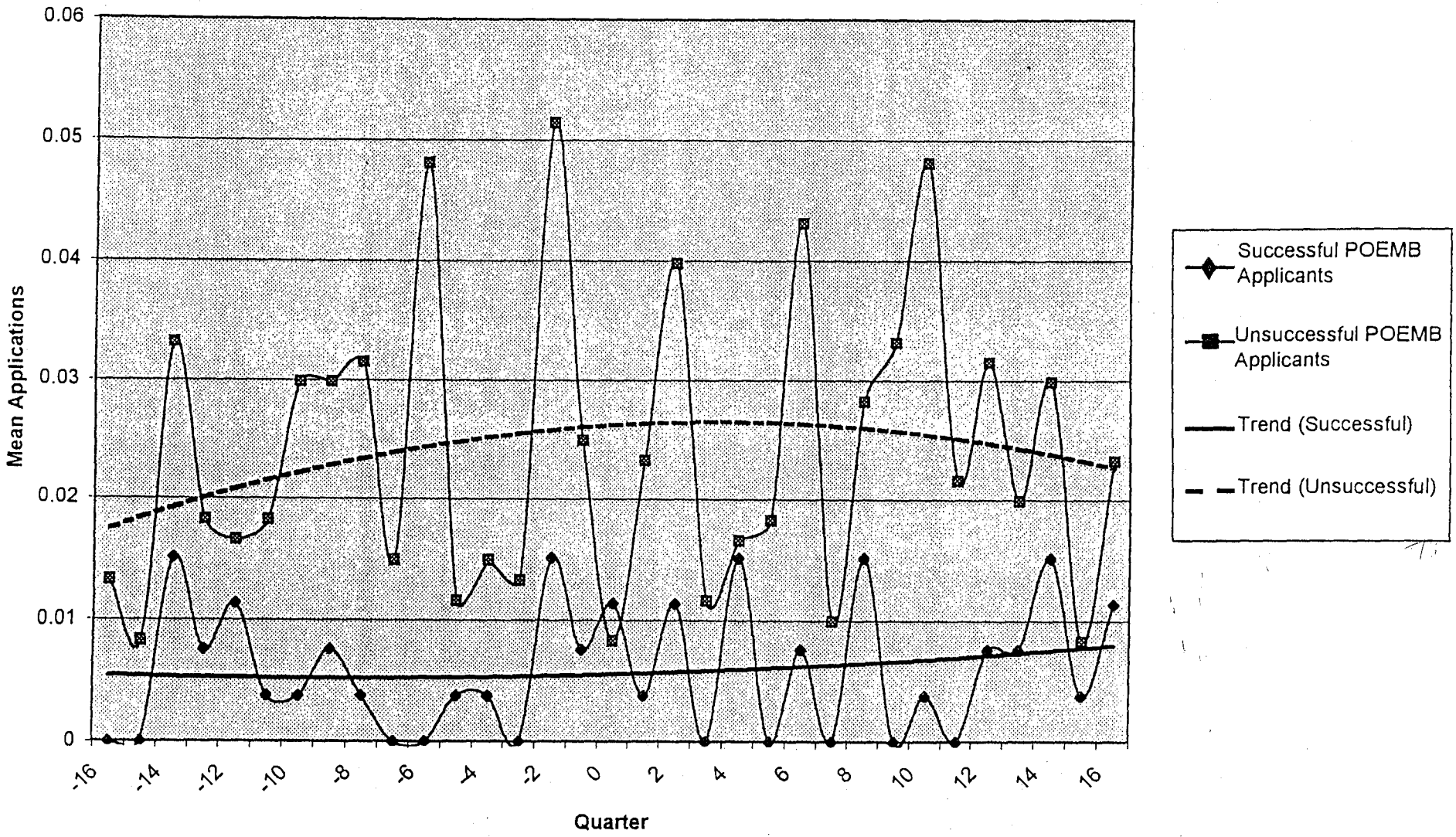


Figure 4.2 Mean Number of Applications, by Quarter (HL Only)



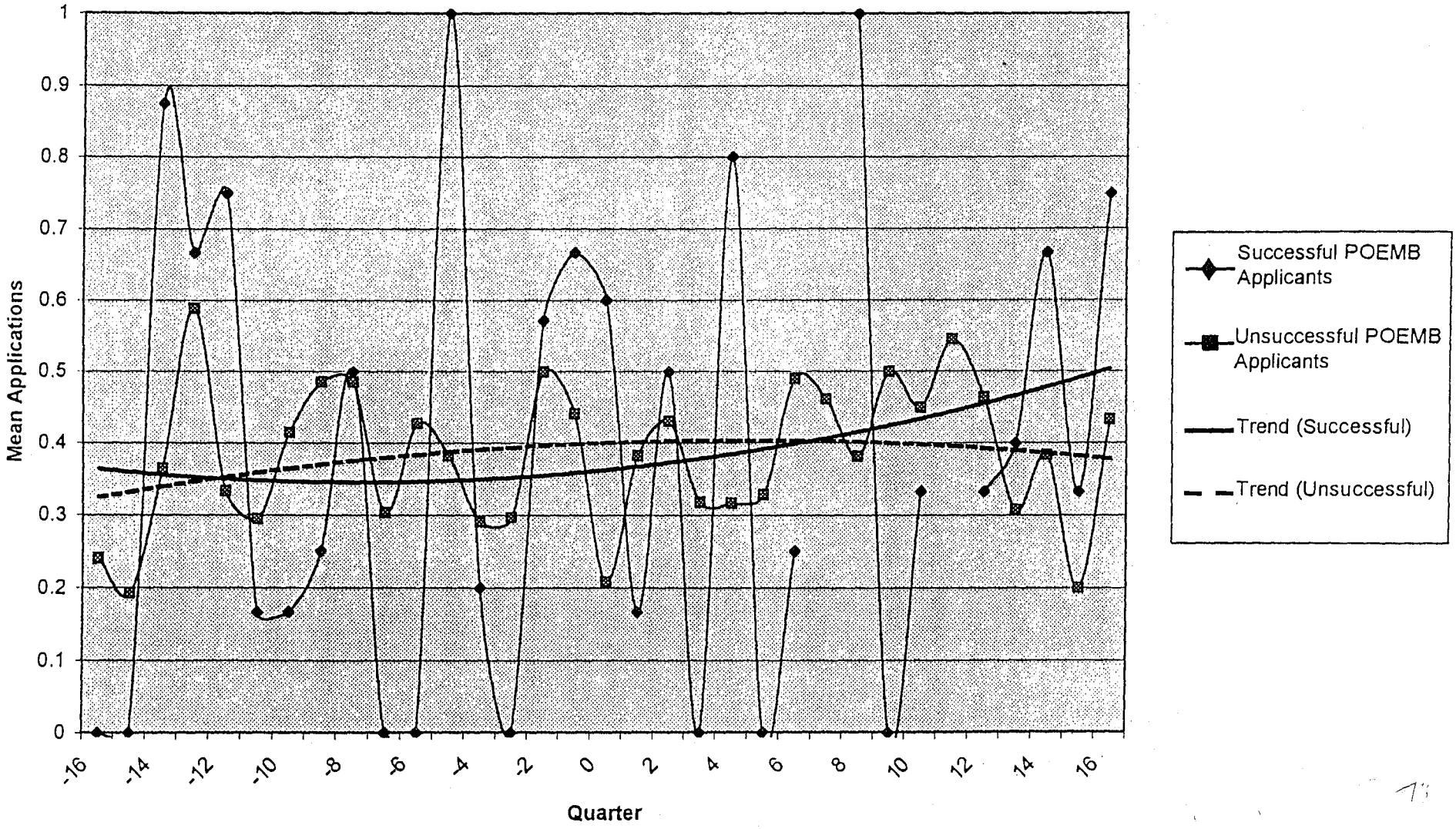
These same data were examined using a mixed-model repeated measures analysis, with POEMB status as a fixed effect, investigator as a random effect nested within POEMB status, the dependent variables specified as polynomial functions of time, and the covariance structure of the repeated observations determined empirically. The fixed main-effect terms included in the model were group (successful POEMB applicants versus unsuccessful applicants), time (in quarters), and time-squared. Time-squared was included so that any curvilinear as well as linear trends could be detected. The interaction terms were group by time and group by time-squared. If receiving a POEMB award affected application activity, we would expect to see a group by time and/or group by time-squared interaction (e.g., relative differences between groups emerging during the post-POEMB period). As expected, the models for both total and NHLBI applications showed main effects of group ($p < 0.01$), no main effects of time, and no group by time interactions in the linear or quadratic terms. Consequently, they provide no evidence consistent with an effect of POEMB on application activity at the investigator level.

In order to determine the impact of POEMB on the research direction/focus of both applicant groups, we also examined the ratio of NHLBI (HL) to total NIH applications for those investigators submitting one or more NIH applications. This allows us to see changes over time in the *proportion* of application activity that is HL-related, holding *total* application activity constant. The longitudinal data and their trends are plotted in Figure 4.3. As shown, the unsuccessful POEMB applicant group's trend is practically flat, while the successful applicant group's ratio starts lower and ends higher than that of the unsuccessful applicant group — a classic group-by-time interaction. Correspondingly, simple before-after comparisons that aggregated data across quarters showed the unsuccessful POEMB applicant group with a higher average ratio of HL to total applications in the pre-award period (41 percent for unsuccessful to 32 percent for successful), while the successful applicant group had a higher average ratio of HL to total applications in the post-award period (49 percent for successful to 33 percent for unsuccessful).[4] This interaction was eliminated when the data were examined by status, however. Once

---

[4] It should be noted, however, that the individuals contributing scores in the post period are not in all cases the same individuals contributing scores in the pre period. This is because the NHLBI/total ratio can only be computed for individuals who submit at least one grant application, and the set of investigators submitting one or more grants in the post period is not identical to the set of investigators submitting one or more grant applications in the pre period. Thus, the ratio data are not strictly longitudinal like the count data. Similarly, a longitudinal analysis could not be conducted on the quarterly data due to the high numbers of missing values in each quarter (i.e., NHLBI/total ratios were only computable in quarters where the investigator submitted one or more applications). This will be done for subsequent analyses of rates, ratios, or proportions.

Figure 4.3 Mean Number of Applications, by Quarter (HL/Total Ratio)



stratified, there was no difference between the ratios of the senior investigators (SCODES 1 and 2) in the two applicant groups. However, when the trainees from the two applicant groups were compared, a different pattern was apparent. There was an increase in the HL/total applications ratio from the pre- to post-POEMB periods for the POEMB trainees and potential trainees (status 4 and 3, respectively), while there was a decrease for the potential trainees (status 3) of the unsuccessful applicant group for the same period. This difference is not statistically significant due to the small *n*, but it does suggest that the higher average ratio of HL to total applications observed in the post-POEMB period for the successful applicants (all statuses) was due to the grant application activity of the POEMB trainees[5]. Overall, this lends support to the hypothesis that receiving the POEMB frees senior investigators from having to apply for multiple grants to maintain funding streams, while stimulating trainees-level researchers to pursue more grants, particularly those designated as HL.

4.1.2.2 Analysis of "Switching"

An additional approach to determining the effect of POEMB on the direction/focus of the research of individuals in both applicant groups is to use the ICD code associated with the grant application to examine the NIH/PHS grant application data in terms of switching from pre-POEMB states in which heart/lung-designated (HL) grant applications were not submitted to ones in which they were. This approach has the advantage that all of the "switches" of interest can be examined.

Tables 4.2 and 4.3 show the complete transition matrices for successful and unsuccessful POEMB applicants, respectively. The transition matrix shows the probability of all possible transitions from pre-POEMB to post-POEMB status. There are four possible pre- and post-states:

- None - no applications submitted to NIH,
- Non-HL-only - one or more applications submitted to NIH, but none submitted to NHLBI,
- Mixed - multiple applications submitted to NIH, at least one of which was submitted to

---

[5] A comparison of the average HL/total applications ratio of the status 3 and 4 personnel shows that, for the total and post-POEMB period, the status 3s exhibit significantly higher ratios ($p < 0.01$) and are thus contributing the most to this ratio. An "apples-to-apples" comparison of status 3 trainees across applicant groups would show a still stronger association between POEMB and the HL/total application ratio among trainees, though the fact that status 3s are not cited in progress reports muddies the interpretation.

NHLBI and one of which was not, and

- HL-only - one or more applications submitted to NHLBI, no other applications submitted to NIH.

Each pre-POEMB state has a separate set of transition probabilities that sum to 100 percent. In Table 4.2, for example, among investigators in the successful POEMB applicant group who submitted no applications pre-award, 88 percent submitted no applications post-award, 4 percent submitted only non-HL applications, 1 percent submitted both HL and non-HL applications, and 7 percent submitted only HL applications.

The shaded cells represent the four transition states that could be categorized as switching, so designated because each represents a transition from a state in which HL grants were not applied for to a state in which they were. These include: none to HL-only, none to mixed, non-HL-only to HL-only, and

Table 4.2 Pre-POEMB Application to Post-POEMB Application Transition Matrix for NIH/PHS Grant Applications Submitted by Successful POEMB Applicants

Pre-POEMB Application	Post-POEMB Application Frequency, Row%				
	no applications	non-HL only	HL only	mixed (HL/non-HL)	Total Investigators
no applications	201 88%	10 4%	15 7%	3 1%	229
non-HL only	8 38%	9 43%	4 19%	0 0%	21
HL only	5 56%	1 11%	3 33%	0 0%	9
mixed (HL/non-HL)	3 60%	2 40%	0 0%	0 0%	5
Total Percent	217 83%	22 8%	22 8%	3 1%	264

Table 4.3 Pre-POEMB Application to Post-POEMB Application Transition Matrix for NIH/PHS Grant Applications Submitted by Unsuccessful POEMB Applicants

Pre-POEMB Application	Post-POEMB Application Frequency, Row%				
	no applications	non-HL only	HL only	mixed (HL/non-HL)	Total Investigators
no applications	189 65%	47 16%	43 15%	13 4%	292
non-HL only	52 33%	80 50%	8 5%	19 41%	159
HL only	37 35%	9 8%	50 47%	10 9%	106
mixed (HL/non-HL)	5 11%	19 41%	12 26%	10 22%	46
Total Percent	283 47%	155 26%	113 19%	52 8%	603

non-HL-only to mixed. As shown in Table 4.3, unsuccessful POEMB applicants with no pre-award applications had a higher probability of transitioning to both mixed (4 percent to 1 percent) and HL-only (15 percent to 7 percent). Unsuccessful POEMB applicants with non-HL-only pre-award applications had a higher probability of transitioning to mixed status (12 percent to 0 percent), while successful POEMB applicants had a higher probability of transitioning to HL-only (19 percent to 5 percent).

Aggregating across the four switching categories, unsuccessful POEMB applicants were more likely to switch (14 percent to 8 percent), and this difference was statistically significant ($p < 0.05$). Unlike the transition matrices, however, the aggregated analysis is unconditional; that is, it fails to consider the differences between groups in their potential to switch. Specifically, if one group is already submitting more HL applications during the pre-award period relative to the other group, then a smaller fraction of its members have the potential to switch. For this reason, we separately examined the subset who were not submitting HL applications during the pre-award period. Seven hundred and one (701) investigators met this criterion: 105 successful POEMB applicants and 596 unsuccessful POEMB applicants. Unsuccessful POEMB applicants were twice as likely to switch (18 percent to 9 percent), and again the difference was statistically significant ($p < 0.01$).

In the above analysis, investigators with and without pre-POEMB application activity are grouped together. It is arguable that those who submit no applications in the pre-POEMB period are not

truly switching when they submit applications in the post-POEMB period; rather they are simply becoming active applicants. Therefore, it is also of potential interest to examine the subset of investigators who are active in the pre-POEMB period. Three hundred and forty-six (346) investigators were in this subset: 31 successful POEMB applicants and 315 unsuccessful POEMB applicants. In contrast to the full sample, successful POEMB applicants were more likely to switch, although the difference was small (11 percent versus 9 percent) and not statistically significant. Finally, we examined the subset of investigators who were active in the pre-POEMB period *and* capable of switching. This subset included 180 investigators (21 successful POEMB applicants and 149 unsuccessful POEMB applicants). Again, successful POEMB applicants were slightly more likely to switch (19 percent to 17 percent), though the difference was not significant.

In sum, there were several switch patterns in which unsuccessful POEMB applicants transitioned at significantly higher rates than did successful applicants, others in which there was no significant difference, and one in which successful POEMB applicants transitioned at significantly higher rates than applicants in the unsuccessful applicant group. Specifically, successful POEMB applicants submitting only non-HL applications in the pre-award period had a higher probability of submitting only HL applications in the post-award period. This finding is consistent with the hypothesis that receiving the POEMB may have led to a shift or switch in focus from nonHL research to HL research as measured by grant application activity and ICD. These switching data offer no other evidence that receiving a POEMB grant increased the probability of investigator switches of the sort hypothesized, at least as measured by the application for NIH/PHS grants.

4.2 Awards — National Institutes of Health, Public Health Service

This section focuses on an examination of NIH/PHS award activity of individuals in the successful and the unsuccessful POEMB applicant groups and is subdivided into sections on (1) methods and (2) findings. The main goal of this examination is to describe the activity in the two applicant groups and to assess the effect (if any) of the POEMB on this activity.

4.2.1 Methods

CRISP descriptors (or subject terms) were used to classify the nature of the NIH/PHS awards received by successful and unsuccessful POEMB applicants during the period 17 quarters before and 16 quarters after POEMB application. First, a list was compiled of all the **CRISP** subject terms associated with the awards received by individuals in both POEMB applicant groups. This list was alphabetized and manually checked using the *CRISP Intramural Research Index* to verify the classification of terms. This index contains a complete listing of **CRISP** subject headings used to describe research projects funded by NIH. Awards with **CRISP** descriptors listed under the major headings "cardiac", "cardiovascular", "heart", and "respiratory" were classified as heart/lung (HL). Awards with descriptors listed under the major heading "molecular biology" were classified as molecular biology (MB). Awards with all other descriptors were classified as "other" (CC). A complete list of **CRISP** descriptors that fall within the categories of interest are presented in Appendix 2A. Each award is associated with one or more descriptors; thus there are more than three possible **CRISP** classification categories. These categories include: HL only, MB only, HL/MB (mixed), "other" only, HL/ "other", MB/ "other", and HL/MB/ "other". This allows for a more sensitive measure of area of research focus than simply focusing on ICD.

The analysis of NIH/PHS award activity was carried out in a manner very similar to that in the previous section. The dependent variables obtained directly from the **CGAF/CRISP** baseline file were (1) total awards, (2) HL awards (**CGAF**), and (3) the distribution of HL, MB, HL/MB, and "other"-designated awards (**CRISP**). Those dependent variables that were constructed using information from the baseline file included (4) the ratios of **CRISP** descriptor categories to total applications, (5) ratios of these categories to total awards, (6) having received one or more total awards versus no awards, (7) having received one or more HL-designated awards versus no awards (**CGAF**), and (8) having received one or more HL, MB, and/or HL/MB-designated awards (**CRISP**).

The analyses focused mainly on the following measures of the differences in the award activity of researchers in the two POEMB applicant groups:

- Aggregate number of total number of awards in the pre-POEMB award, post-POEMB award, and total periods.

- Mean number of total, HL, MB, HL/MB, and "other"-designated awards in the pre-POEMB award, post-POEMB award, and total periods.
- Repeated-measures analysis of the above-noted dependent variables.
- Success rate (total awards to total applications) and HL, MB, HL/MB, and "other"-designated awards as a proportion of total applications in the pre-POEMB award, post-POEMB award, and total periods.
- HL, MB, HL/MB, and "other"-designated awards as a proportion of total awards in the pre-POEMB award, post-POEMB award, and total periods.
- Visual presentation of each dependent variable over time (17 quarters pre-POEMB award to 16 quarters post-POEMB award), with trend lines to fit the data.
- Post-POEMB differences in **CRISP** descriptor distributions versus pre-POEMB differences in these distributions (Chi-Square test).
- Transition probabilities from all possible pre-POEMB states to all possible post-POEMB states.
- Proportional differences in terms of "switches" (changes from certain **CRISP** descriptor categories in the pre-POEMB period to other key categories in the post-POEMB period).

These differences were examined by pooling all status types (SCODEs 1, 2, 3, and 4), and then comparisons were made between the senior investigators (SCODEs 1 and 2) in the two POEMB applicant groups as well as between trainee-level personnel (SCODEs 3 and 4 for successful applicants are compared to SCODE 3 for unsuccessful applicants). As with applications, it was important to break out senior investigators and trainees because the program was expected to affect senior investigators and trainees differently on criterion dependent variables, and also because the percentage of senior investigators was far lower in the successful than in the unsuccessful applicant group, leading to potentially misleading results when the two subgroups are pooled.

As in the previous section, visual comparisons over the entire pre-post period (longitudinal) are an important focus of this analysis. These comparisons are provided in the form of plots over time and trend lines to fit plots.

The (null) hypothesis tested in this analysis is that there are no (statistically significant) group by time interactions between the two POEMB applicant groups in terms of the measures listed above. Given

the nature of the POEMB, we would expect that receiving the POEMB would have a "shielding" effect on the senior investigators (SCODEs 1 and 2), in other words, these investigators would be freed from having to submit multiple applications to maintain a stream of funding for their research, which would presumably allow them to spend more of their time on research. This effect may manifest itself as a decrease in the total awards received by the POEMB senior investigators during the post-POEMB application period. The research activity of POEMB trainees and potential trainees (SCODEs 4 and 3, respectively), however, may have been stimulated by the program, leading them to pursue funding (submit grant applications and receive awards) for more new research projects. We would also expect that for the successful applicants, receiving the POEMB may have led to one of two possible pre- to post-POEMB shifts in research focus as measured by **CRISP** descriptor categories — a shift from a focus that is primarily on cardiovascular or pulmonary disease research to one that incorporates molecular biologic techniques in the study of these diseases, or a shift from a focus that is primarily on the application of molecular biologic techniques to nonheart/lung related studies to one that applies those techniques to fundamental cardiovascular or pulmonary research problems. It is conceivable that similar effects might be observed among unsuccessful applicants as well, as a "ripple effect" of applying for the POEMB.

In all cases where the data support it, statistical modeling was carried out on the full sample (senior investigators and trainees pooled) in the form of repeated measures analysis. As with the application data, all such analyses were conducted using PROC MIXED in SAS®, with POEMB as a fixed effect, investigator as a random effect nested within POEMB, the dependent variables (separately) as polynomial functions of time (17 quarters before to 16 quarters after POEMB application), and the covariance structure of repeated observations determined empirically (compound symmetry or autoregressive).

4.2.2 Findings

Below we discuss findings related to **CRISP** descriptors as a measure of the subject matter of NIH/PHS awards in terms of (1) longitudinal analysis and (2) "switching" analysis.

4.2.3.1 Longitudinal Analysis

Considering all statuses (SCODEs) together, the successful POEMB applicant group was successful on 59 percent of the applications they submitted to NIH/PHS during the period 17 quarters before and 16 quarters after POEMB application compared to a 48 percent success rate for the unsuccessful POEMB applicant group for the same time period. Table 4.4 shows the total number of NIH/PHS applications that were funded for both POEMB applicant groups by status.

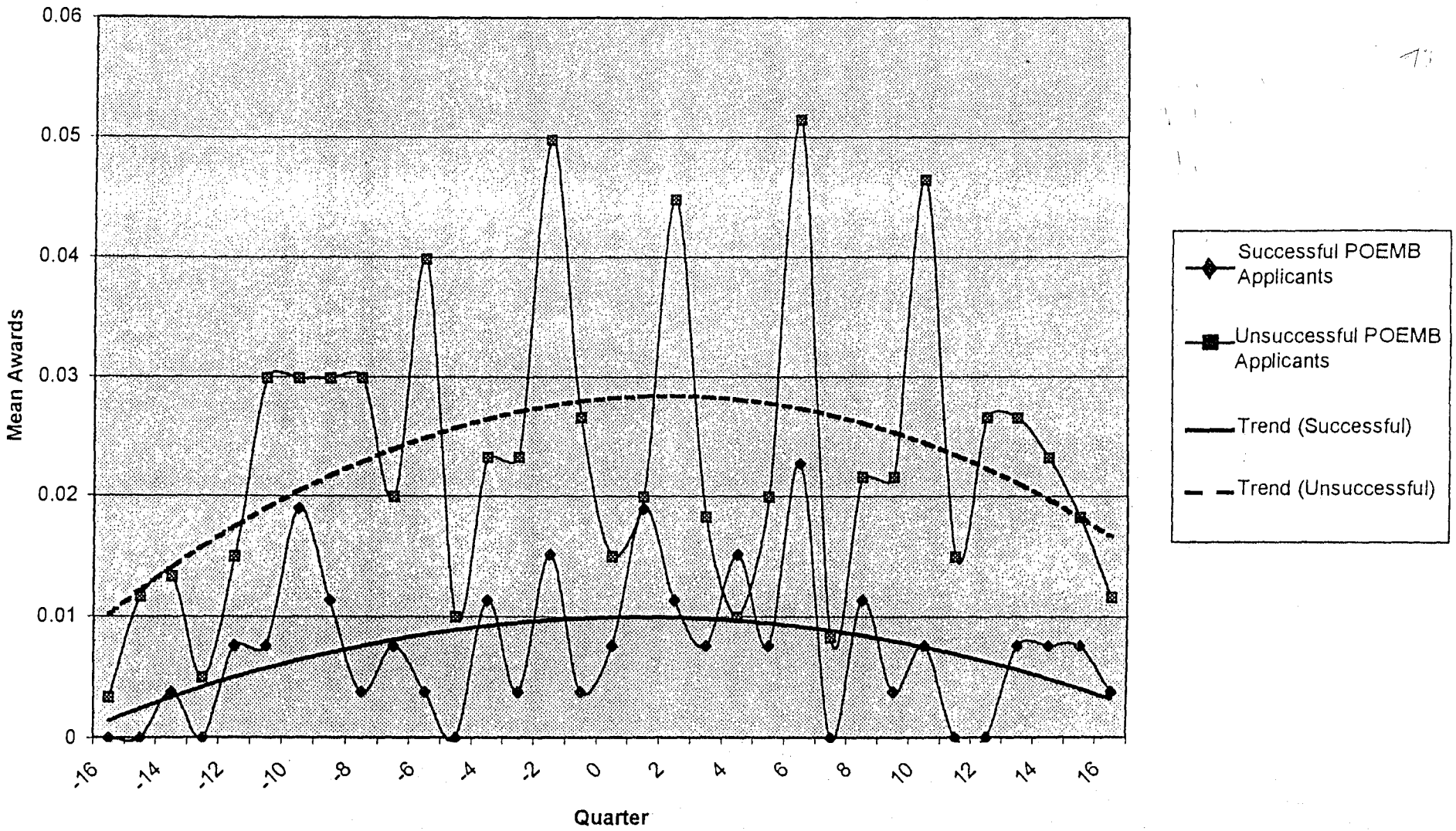
Total Awards. Figure 4.4 shows the average number of total awards over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each time series. The trends indicate greater numbers of awards for unsuccessful POEMB applicants throughout the pre- and post-award periods, consistent with their greater levels of application activity. Like the application data, the award data were analyzed using the repeated measures mixed model with linear and quadratic terms. That analysis showed main effects of program ($p<0.01$) and time-squared ($p<0.01$), but no interaction effect of program by time in the linear or quadratic terms. (The quadratic effect of time is visible in the curvilinear trends of Figure 4.4.) That is, like the analysis of applications, the data provide no evidence consistent with an effect of POEMB on total awards at the investigator level.

Simple before-after comparisons of the average number of total awards showed three-fold or greater differences between groups for the aggregate, pre-POEMB, and post-POEMB periods (all significant at $p<0.01$). There was some narrowing of the mean difference across periods, but it was small (less than 0.02 awards/investigator) and statistically nonsignificant. Examination of these same data by status code (SCODE) eliminated the three-fold or greater differences in awards between the senior investigators (SCODEs 1 and 2) in the successful and unsuccessful POEMB applicant groups for all three periods (total, pre and post), and none of the differences were statistically significant. That is, the three-fold difference in the pooled data was driven by the far higher percentage of trainees in the successful applicant group (87 percent versus 37 percent), as described in the methods section. A comparison of the total number of awards received by the status 3 and 4 POEMB trainee-level researchers shows that, for the total and post-POEMB period, the status 4s (trainees cited in POEMB progress reports) received significantly more awards than the status 3s or potential trainees ($p<0.001$). Had the status 4s been removed from the original analysis, the successful POEMB applicants would have exhibited an even lower total number of awards relative to the unsuccessful POEMB applicant group

Table 4.4 Grant Applications, Funded and not Funded, by Investigator Status

Investigators from Successful POEMB Applicant Group	Applications Submitted 16 Quarters Pre-POEMB Frequency, Row %, Column%			Applications Submitted 16 Quarter Post-POEMB Frequency, Row %, Column%			Total
	Funded	Not funded	Total	Funded	Not funded	Total	
Status 1 and 2	31	18	49	22	11	33	82
(Status 1 N=4; Status 2 N=30)	63%	37%		67%	33%		65%
	82%	72%		61%	41%		
Status 3	2	2	4	3	2	5	9
(N=32)	50%	50%		60%	40%		7%
	5%	8%		8%	7%		
Status 4	5	5	10	11	14	25	35
(N=198)	50%	50%		44%	56%		28%
	13%	20%		31%	52%		
Investigators from Unsuccessful POEMB Applicant Group							**Total**
Status 1 and 2	242	284	526	215	297	512	1038
(Status 1 N=18; Status 2	46%	54%		42%	58%		84%
N=365)	82%	90%		73%	91%		
Status 3	54	30	84	79	31	110	194
(N=220)	64%	36%		72%	28%		16%
	18%	10%		27%	9%		

Figure 4.4 Mean Total Awards Per Investigator, By Quarter



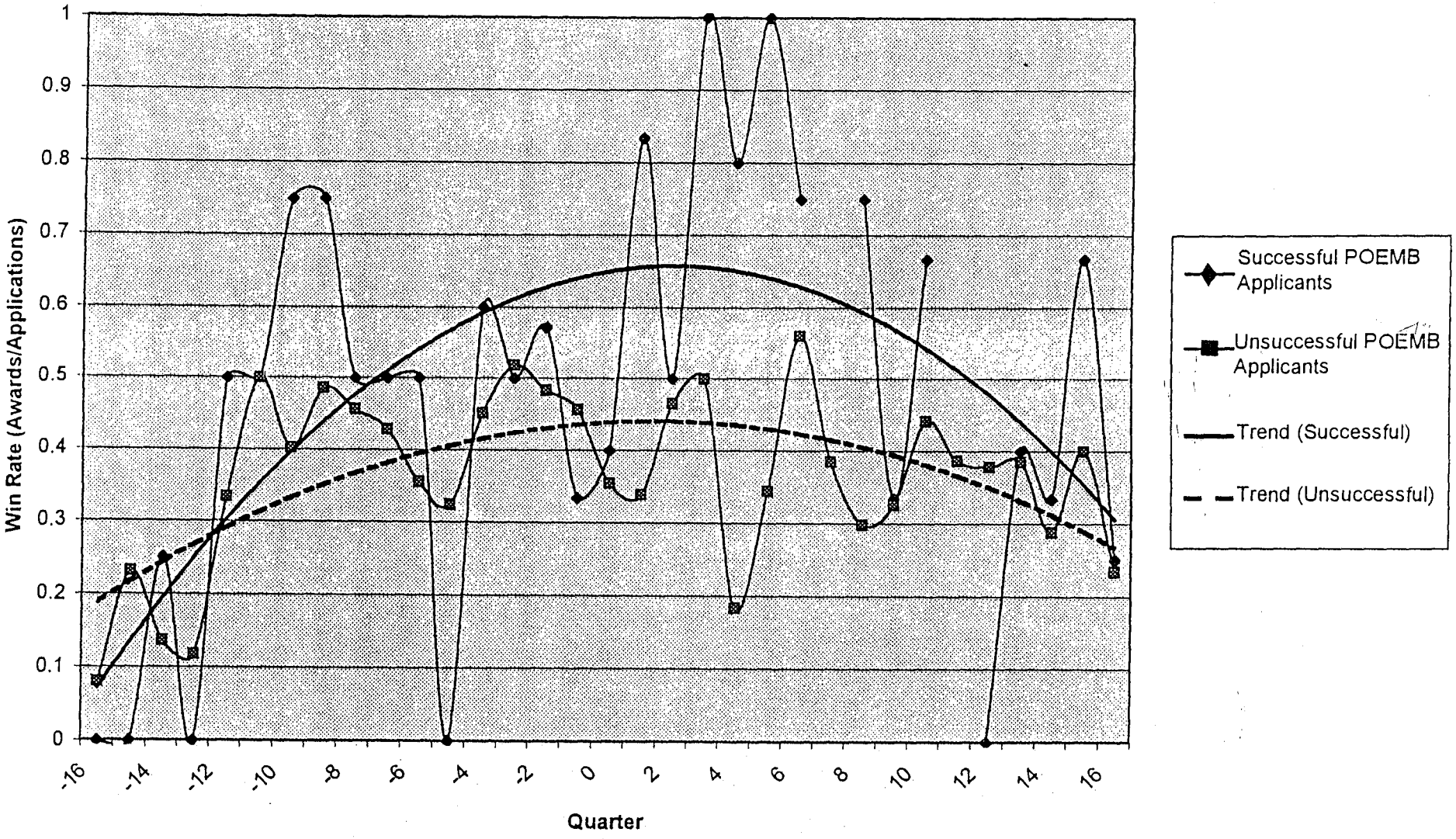
during the post-POEMB period. Thus, the inclusion of these individuals did not bias the results against the POEMB applicant group.

Success Rate. Given that unsuccessful POEMB applicants submitted so many more HL applications on average than did successful POEMB applicants, it is not surprising that they also received more awards. This does not rule out the possibility that POEMB influenced the likelihood of receiving an award, given an application, i.e., the success rate. Figure 4.5 shows the average success rate over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each series. The trends indicate that successful POEMB applicants did tend to register higher success rates in the period following the POEMB application date. Simple before-after comparisons showed that among investigators submitting one or more applications, successful POEMB applicants registered significantly higher success rates when aggregated across both pre- and post-award periods (55 percent to 38 percent, $p<0.01$). Consistent with the trends in Figure 4.5, the difference was nonsignificant when isolated to the pre-award period (45 percent to 38 percent), but significant when isolated to the post-award period (55 percent to 40 percent, $p<0.05$), consistent with a possible impact of POEMB. This is shown visually in Figure 4.5.[6] When analyzed by status, however, the difference between the two POEMB groups' success rates was eliminated for the senior investigators (SCODEs 1 and 2). A comparison of the trainees and potential trainees (SCODES 3 and 4 for successful applicants, SCODE 3 for unsuccessful applicants) in the two POEMB applicant groups showed that the statistically significant difference observed in the post-POEMB period for the pooled (all SCODEs together) successful POEMB applicant group is due to the trainee-level personnel in this group receiving more awards during this period. A comparison of the success rates of the status 3 and 4 POEMB trainees shows that, for the total and post-POEMB period, the status 4s (trainees cited in POEMB progress reports) exhibit a significantly higher success rate ($p<0.001$), and are thus contributing the most to this rate. The results of this comparison indicate that including the status 4s in the POEMB "trainee" group increases the apparent effect of the POEMB during the post-POEMB period.

---

[6] As with the analysis of applications, the individuals contributing scores in the post period are not in all cases the same individuals contributing scores in the pre period, because the success rate can only be computed for individuals who submit at least one grant, and the set of investigators submitting one or more grants in the post period is not identical to the set of investigators submitting one or more grants in the pre period.

Figure 4.5 Mean Total Win Rate, By Quarter



Heart and Lung-designated Awards. Figure 4.6 shows the average number of HL-designated awards over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each time series. The trends indicate greater numbers of HL awards for unsuccessful POEMB applicants throughout the pre- and post-award periods, consistent with their greater levels of HL application activity observed in Section 4.2.3.1. As with total awards, the repeated measures analysis showed main effects of program ($p < 0.01$) and time-squared ($p < 0.05$), but no interaction effect of program by time in the linear or quadratic terms. Simple before-after comparisons showed eight-fold or greater differences between groups for the aggregate, pre, and post periods (all significant at $p < 0.01$). There was some narrowing of the mean difference across periods, but it was small (less than 0.02 awards/investigator) and statistically nonsignificant. Examination of these same data by status code (SCODE) eliminated the eight-fold or greater difference in HL awards between the successful and unsuccessful POEMB applicants for all three periods (total, pre and post). Although much reduced, the difference between the average number of total HL awards received by senior investigators (SCODEs 1 and 2) in the two applicant groups remained statistically significant ($p < 0.01$) for the total period, with successful POEMB applicants receiving fewer awards overall.

As with total awards, unsuccessful POEMB applicants submitted many more HL applications on average than did successful POEMB applicants, so it is not surprising that they also received more HL-designated awards. This does not rule out the possibility that POEMB influenced the likelihood of receiving an HL award, given an application. Figure 4.7 shows HL-designated awards as a proportion of total applications over time. This effectively "controls for" the discrepancy in application frequency. The differences between groups are smaller, yet still consistently present throughout the pre- and post-award periods. Simple before-after comparisons showed almost identical differences across the two periods; 10 percent to 5 percent pre-award and 10 percent to 4 percent post-award for unsuccessful and successful POEMB applicants, respectively. Examination of these data by status showed patterns consistent with the results of the analysis in which all statuses were pooled within the applicant groups.

Finally, POEMB could fail to influence the total number of HL awards and the HL success rate, and still influence the portion of investigators' "portfolios" devoted to HL-designated awards, i.e., HL-designated awards as a proportion of total awards. Figure 4.8 shows this indicator over time. Once again, the trends indicate higher proportions for unsuccessful POEMB applicants throughout the pre- and

Figure 4.6 Mean HL Awards Per Investigator, By Quarter

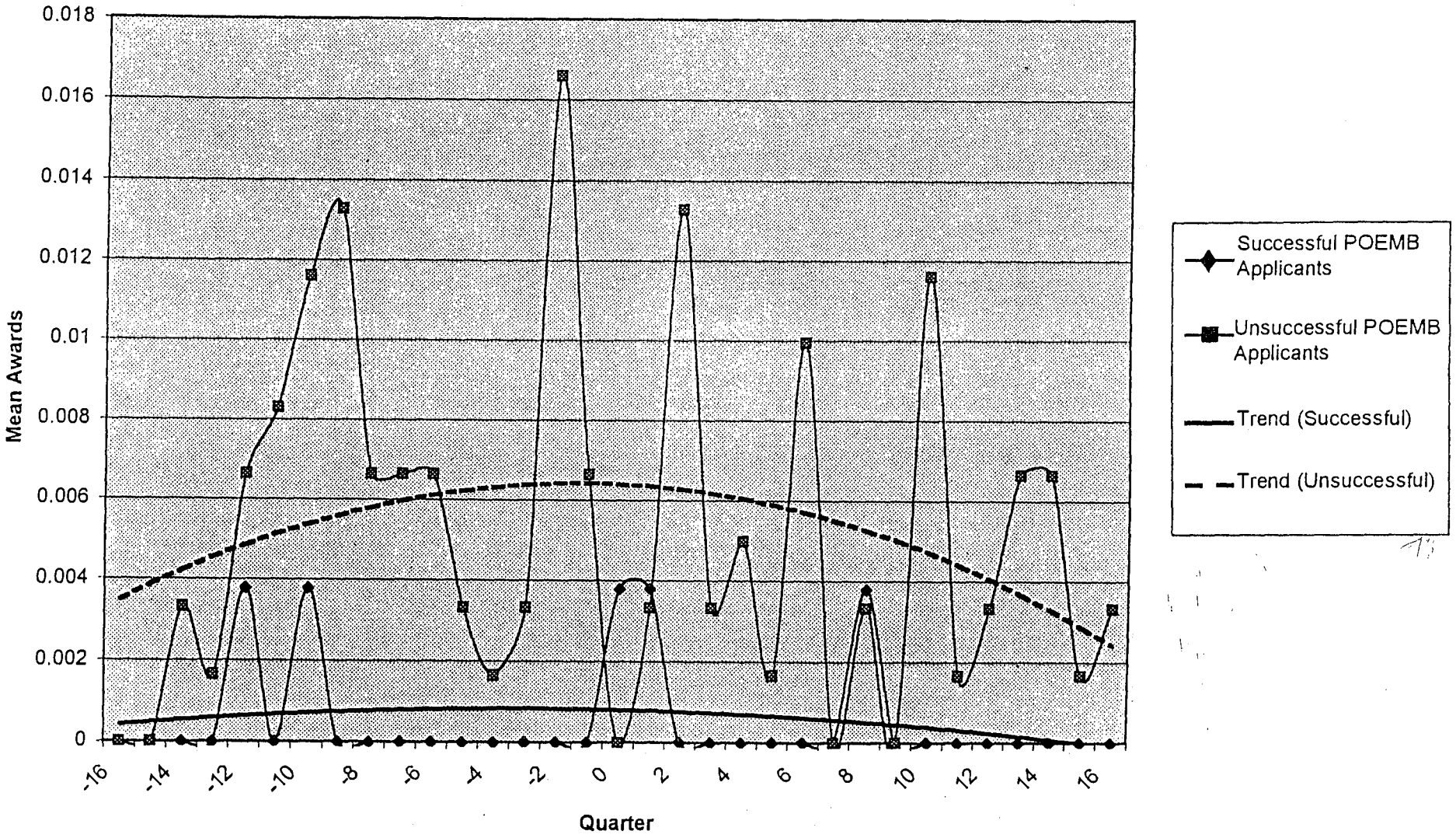


Figure 4.7 Mean Proportion of Applications Yielding HL Awards, By Quarter

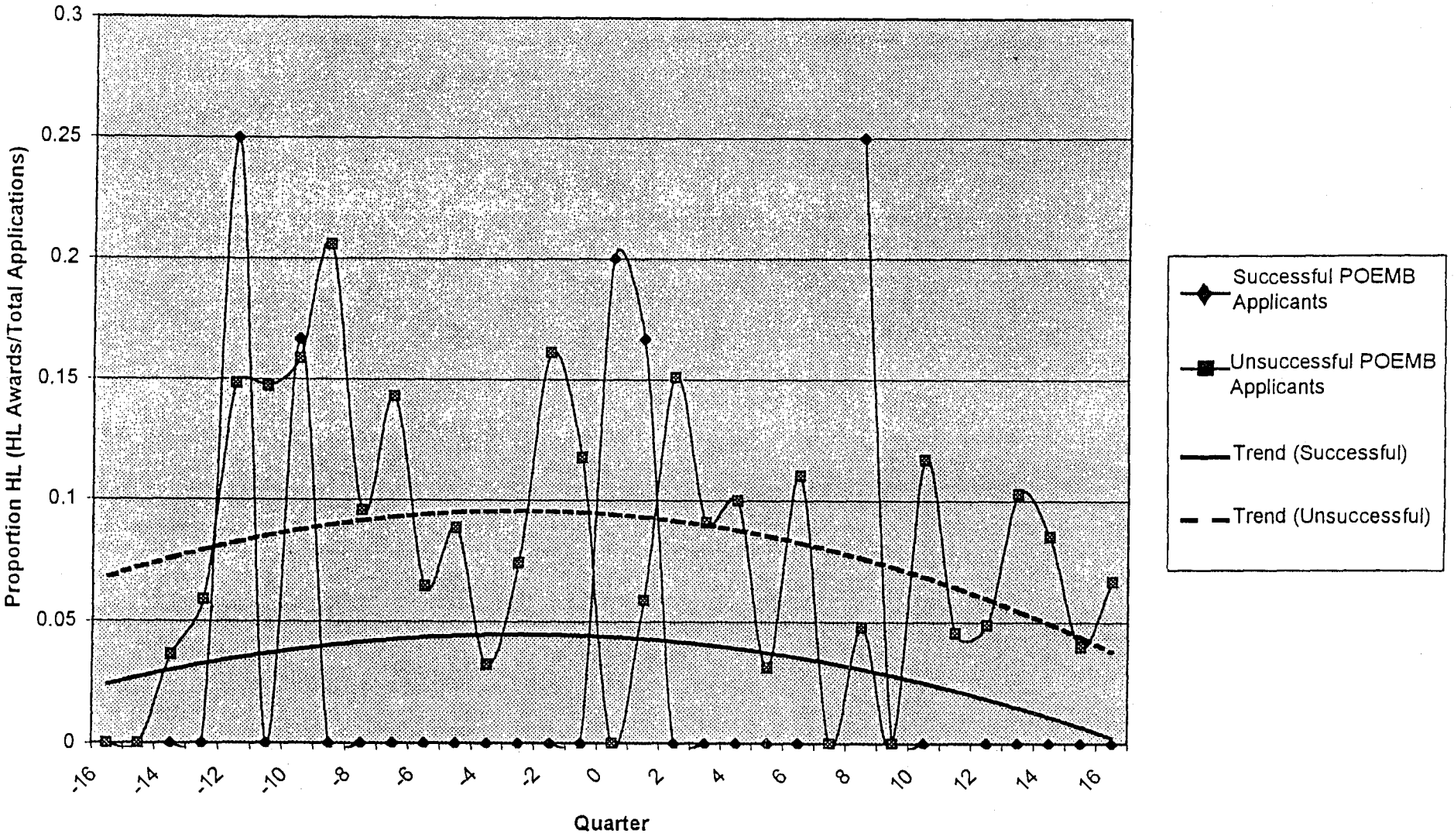
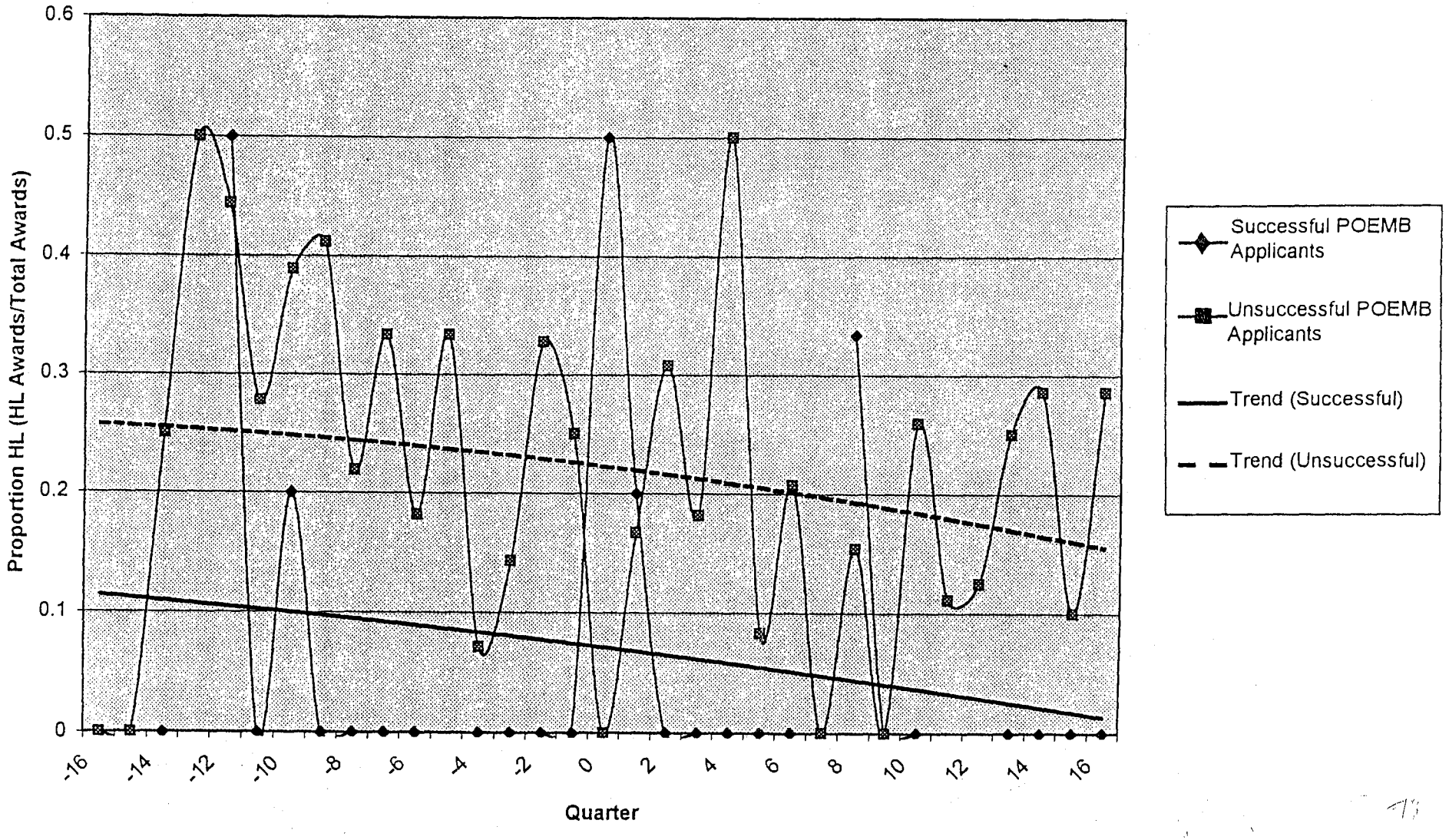


Figure 4.8 Mean Proportion of Total Awards Designated HL, By Quarter



post-award periods. Before-after comparisons showed that among investigators receiving one or more awards, unsuccessful POEMB applicants registered significantly higher rates when aggregated across both pre- and post-award periods (24 percent to 9 percent, $p<0.01$), as well as isolated to the pre-award period (28 percent to 11 percent, $p<0.05$ and post-award period (22 percent to 5 percent, $p<0.01$). Examination of these data by status showed patterns consistent with the results of the pooled status (SCODEs 1, 2, 3 and 4) comparisons discussed above.

Molecular Biology-designated Awards. Figure 4.9 shows the average number of MB-designated awards over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each time series. The trends indicate greater numbers of MB awards for unsuccessful POEMB applicants throughout the pre- and post-award periods. As with total awards and HL-designated awards, the repeated measures analysis showed main effects of program ($p<0.01$) and time squared ($p<0.05$), but no interaction effect of program by time in the linear or quadratic terms. Simple before-after comparisons showed two-fold or greater differences between groups for the aggregate, pre, and post periods (all significant at $p<0.01$). Examination of these data by status showed a nonstatistically significant difference between the average number of MB-designated awards received by senior investigators (SCODEs 1 and 2) in the two POEMB applicant groups. For all periods (total, pre and post), senior investigators from the successful POEMB applicant group received a greater average number of MB-designated awards, with a slight decrease from the pre to the post period.

As with HL awards, the finding that unsuccessful POEMB applicants received more MB awards on average than successful applicants is not surprising given the differences in application frequency. As with HL awards, therefore, we also looked at conditional likelihood of receiving an MB award, given an application. Figure 4.10 shows MB-designated awards as a proportion of total applications over time. Unlike HL awards, the trend lines show successful POEMB applicants with a higher average proportion during most of both periods. Before-after comparisons showed a slightly larger difference in the pre-POEMB period (21 percent to 12 percent) than the post-POEMB period (17 percent to 12 percent). Neither difference was statistically significant. Examination of these data by status showed patterns consistent with the results of the pooled comparisons discussed above. A comparison of the total number of MB-designated awards received by the status 3 and 4 POEMB trainee-level researchers shows that, for the total and post-POEMB period, the status 4s (trainees cited in POEMB progress reports) received

Figure 4.9 Mean MB Awards Per Investigator, By Quarter

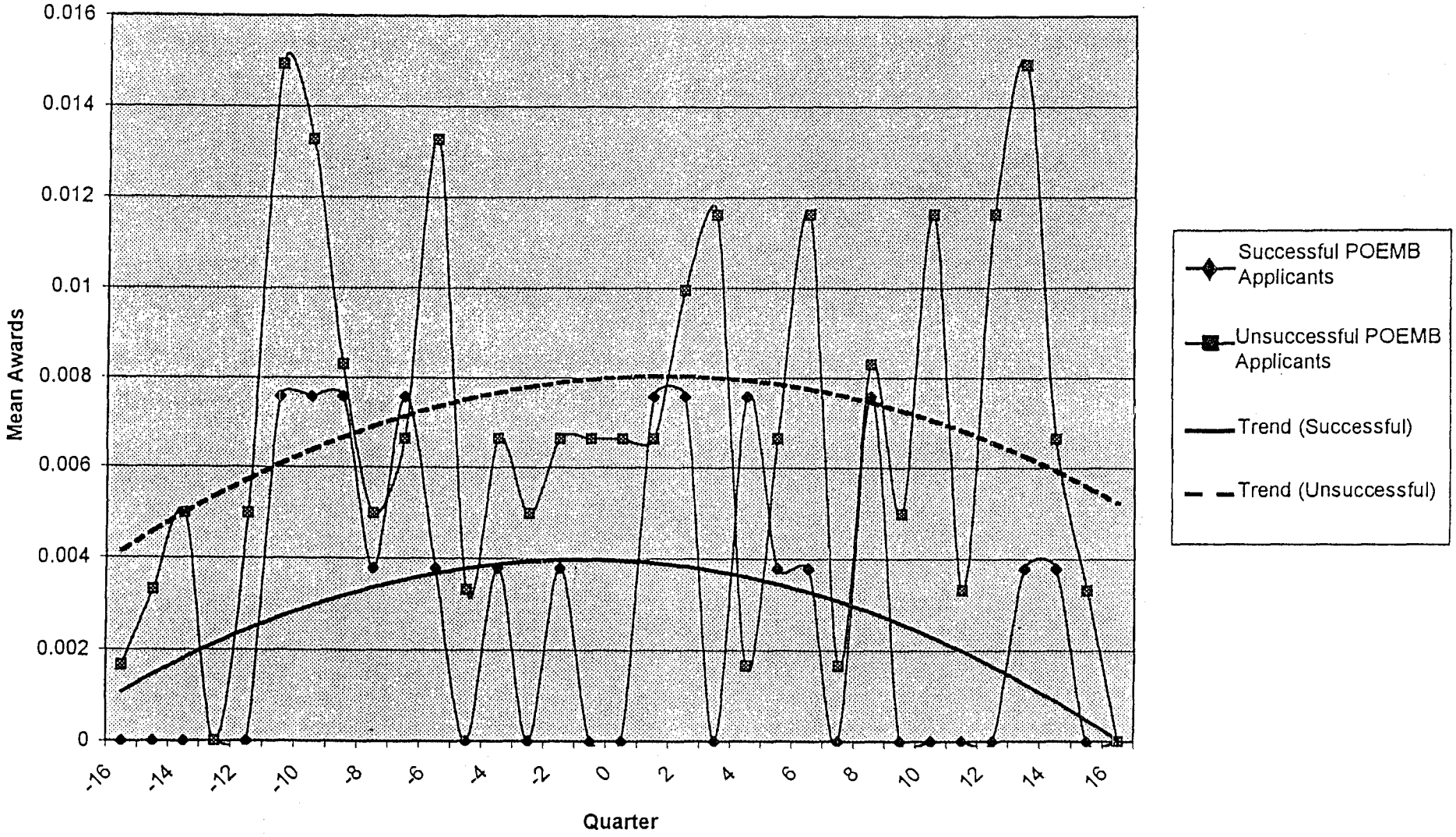
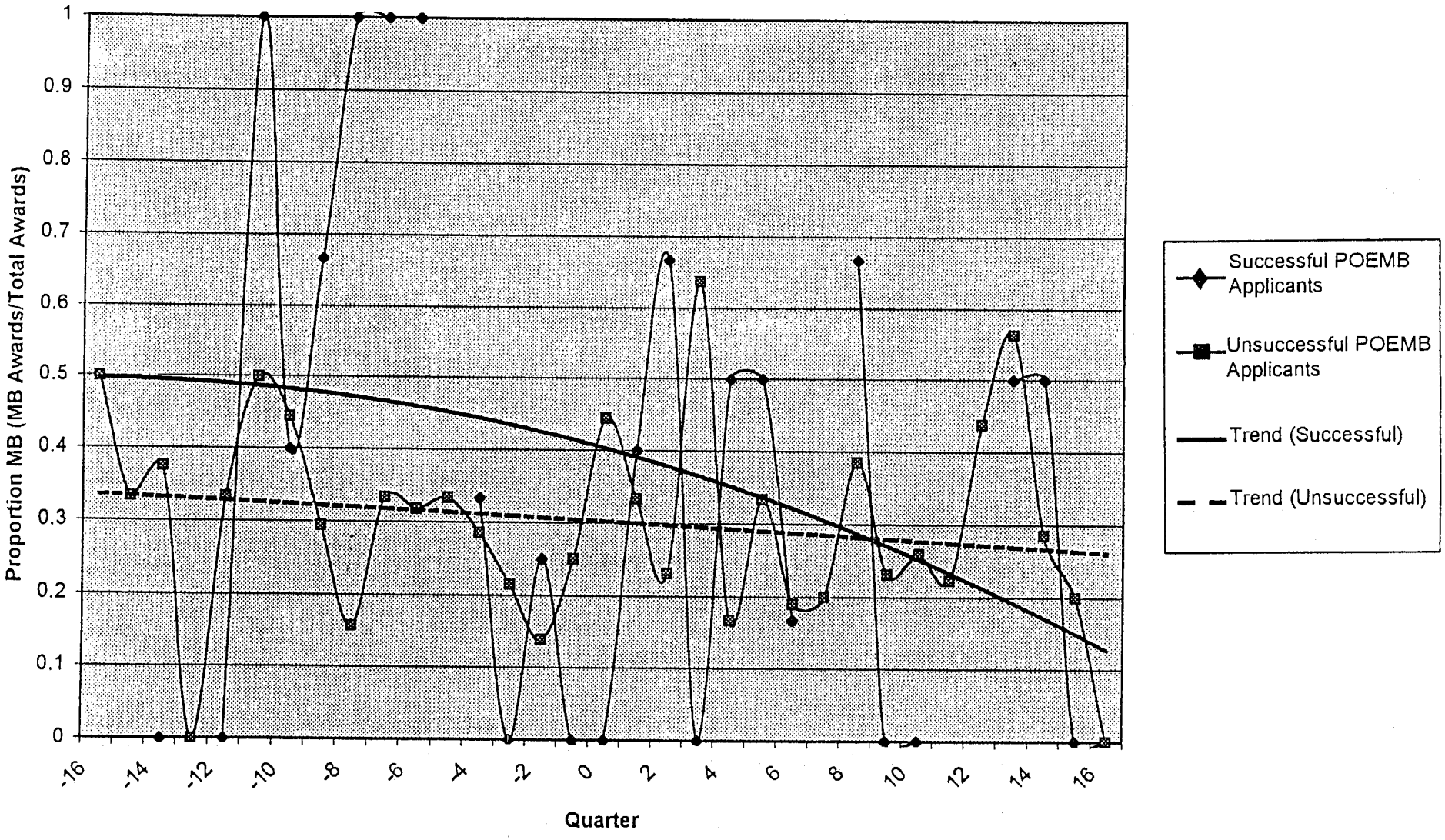


Figure 4.10 Mean Proportion of Total Awards Designated MB, By Quarter



significantly more awards than the potential trainees or status 3s ($p<0.05$). Had the status 4s been removed from the original analysis, the successful POEMB applicants would have exhibited an even lower total number of MB-designated awards relative to the unsuccessful POEMB applicant group during the post-POEMB period. Thus, the inclusion of these individuals did not bias the results against the POEMB applicant group.

Figure 4.11 shows the portion of investigators' "portfolios" devoted to MB-designated awards, i.e., MB-designated awards as a proportion of total awards. Unlike HL awards, the trend lines show successful POEMB applicants with a higher average proportion during the pre-award period that is not sustained through the post-award period. Before-after comparisons confirmed that among investigators receiving one or more awards, successful POEMB applicants registered higher rates of MB awards in the pre-award period (48 percent to 29 percent), though the difference was nonsignificant due to low power (among successful POEMB applicants, only 21 investigators had won NIH awards during this period). The difference completely disappeared in the post-award period (33 percent in both groups). Examination of these data by status showed patterns consistent with the results of the pooled comparisons discussed above. A comparison of the average MB-designated awards/total awards ratios of the status 3 and 4 POEMB trainee-level researchers shows that, for the total and post-POEMB period, the status 4s (trainees cited in POEMB progress reports) received significantly more awards than the potential trainees or status 3s ($p<0.05$). Had the status 4s been removed from the original analysis, the difference between the successful and the unsuccessful POEMB applicants would have been further reduced during the post-POEMB period. Thus, the inclusion of these individuals did not bias the results against the POEMB applicant group. In fact, this finding may suggest that the POEMB did indeed attract researchers interested in applying molecular biologic techniques to study of cardiovascular and/or pulmonary diseases.

Heart and Lung/Molecular Biology-designated Awards. Figure 4.12 shows the average number of HL/MB-designated awards over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each time series. The trends indicate greater numbers of HL/MB awards for unsuccessful POEMB applicants throughout the pre- and post-award periods. The repeated measures analysis showed main effects of program ($p<0.01$), but no main effects of time or interaction effects of program by time in the linear or quadratic terms. Simple before-after

Figure 4.11 Mean Proportion of Applications Yielding MB Awards, By Quarter

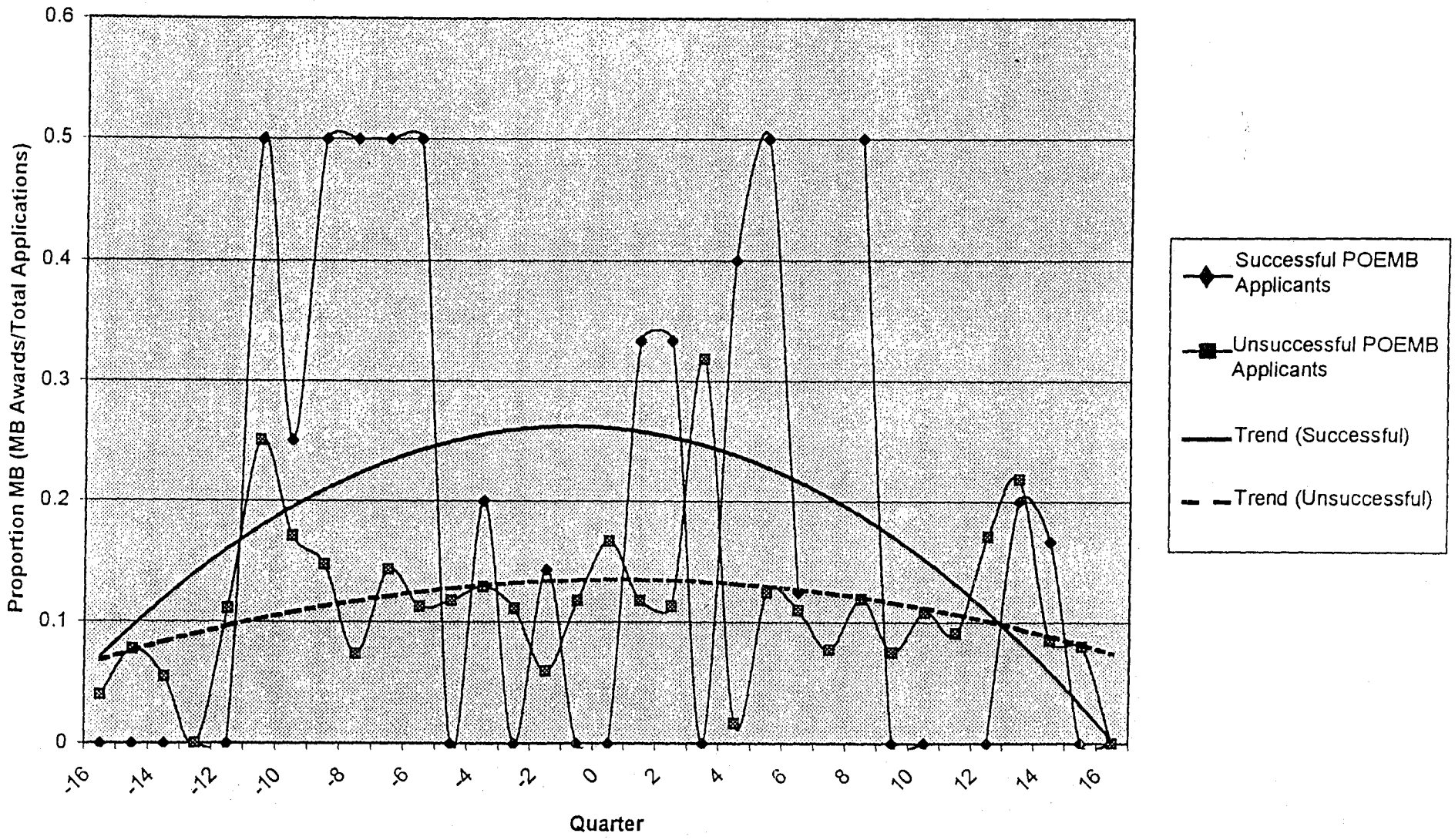
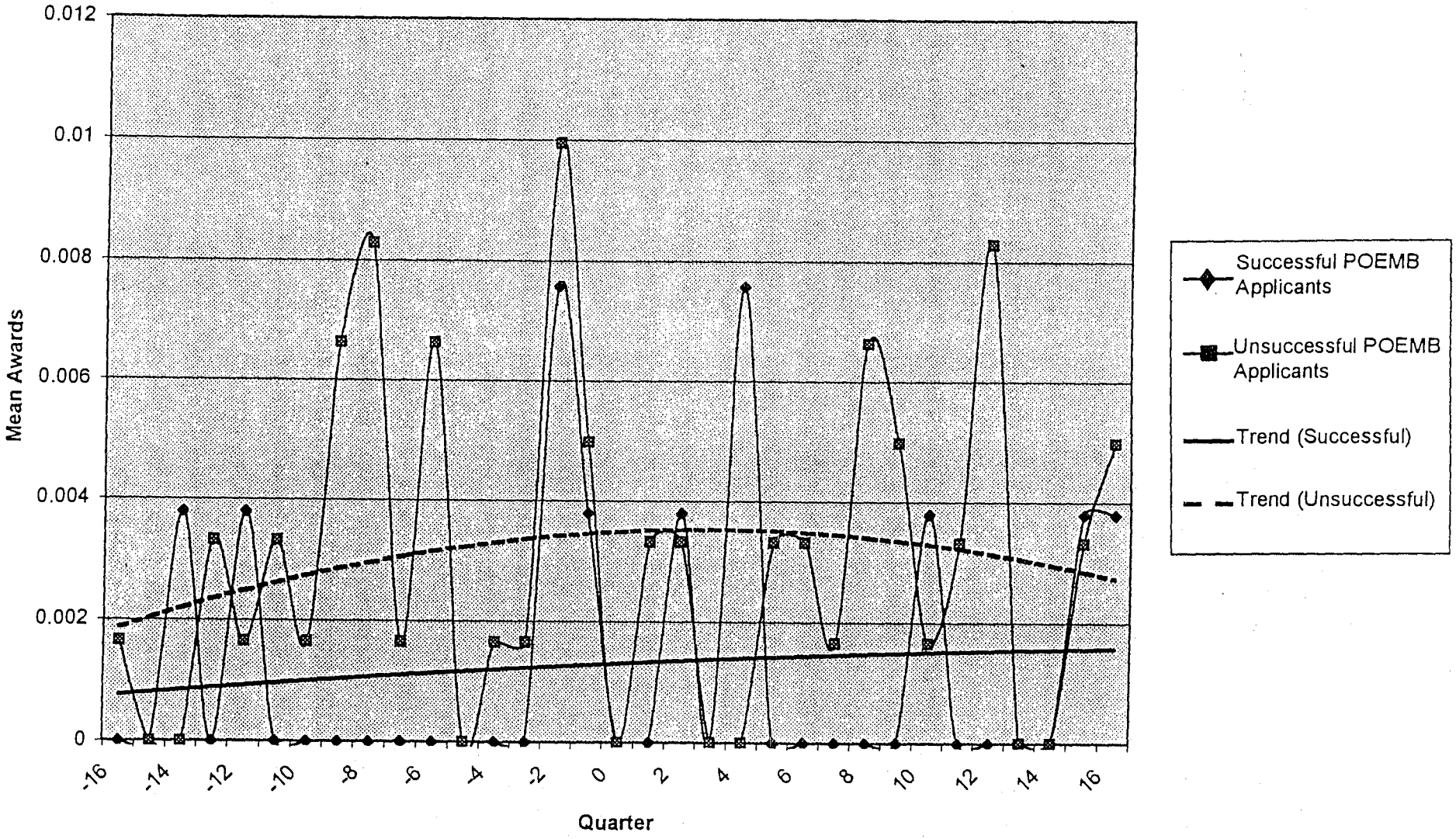


Figure 4.12 Mean HL/MB Awards Per Investigator, By Quarter



comparisons showed two-fold or greater differences between groups for the aggregate ($p < 0.01$), pre ($p < 0.05$), and post periods (nonsignificant due to high variability). The same comparisons by status showed a pattern consistent with the pooled comparisons (SCODEs 1, 2, 3 and 4), with the exception that the average number of HL/MB -designated awards received by senior investigators from the successful applicant group was even greater during the pre-period (in other words, they started higher and then went down). This difference was not statistically significant.

Figure 4.13 shows HL/MB-designated awards as a proportion of total applications over time. While the high variability in the successful POEMB applicant group warrants caution, the pattern of the trend lines is perhaps the clearest yet presented of an effect that is consistent with the hypothesis that receiving a POEMB grant affected investigator behavior. The two trend lines were flat and virtually identical throughout the pre-award period, whereas only the successful POEMB applicants' line increased in the post-award period. Simple before-after comparisons showed a slightly higher rate among unsuccessful POEMB applicants in the pre-award period (5 percent to 4 percent) but a clearly higher rate among successful POEMB applicants in the post-award period (13 percent to 5 percent, nonsignificant difference due to high variability). Examination of these data by status eliminated this difference in rates for all periods for senior investigators. This suggests that it is the awards received by the trainee-level researchers (SCODEs 3 and 4) in the successful POEMB applicant group that are affecting the difference in the averages in the pooled pre-post comparison discussed above. Such a pattern is consistent with the hypothesis that the trainees are maturing, striking out on their own, and integrating molecular biology and cardiovascular/pulmonary research in their work.

Figure 4.14 shows HL/MB-designated awards as a proportion of total awards over time. The trends suggest greater average proportions of HL/MB awards for successful POEMB applicants at the two ends of the observation period, with no difference in the period around the POEMB award date. Simple before-after comparisons paint a somewhat different picture, specifically, a crossover pattern with unsuccessful POEMB applicants posting higher proportions in the pre-award period (14 percent to 12 percent) and successful applicants posting higher proportions in the post-award period (21 percent to 14 percent). Neither of these differences, nor the still larger differences in pre-post gain scores among investigators receiving one or more awards in both periods (14 percent to -2 percent for successful and unsuccessful POEMB applicants, respectively) were statistically significant due to small *n's* in the successful applicant group. A simple before-after comparison by status showed that the proportions of

Figure 4.13 Mean Proportion of Applications Yielding HL/MB Awards, By Quarter

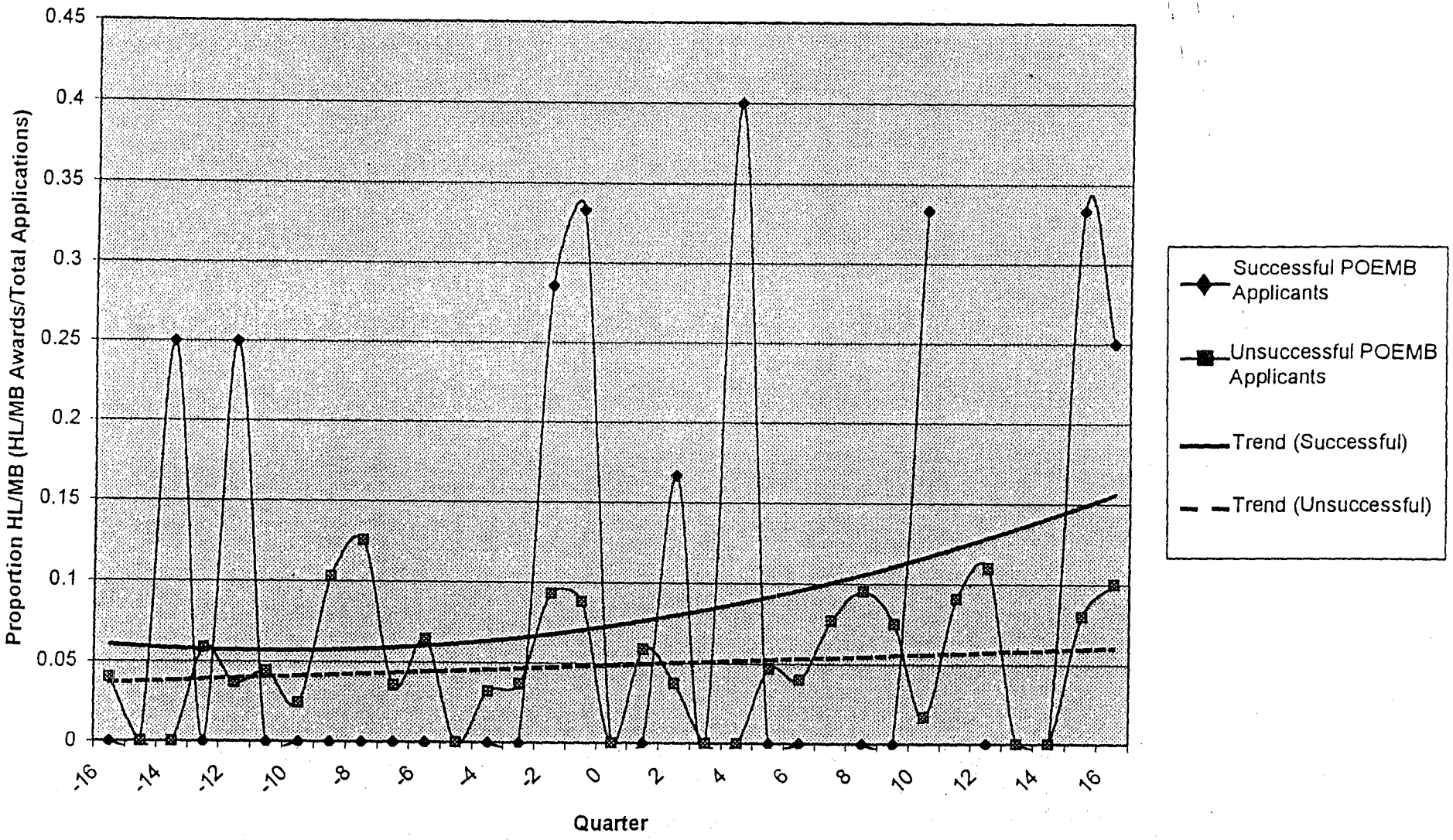
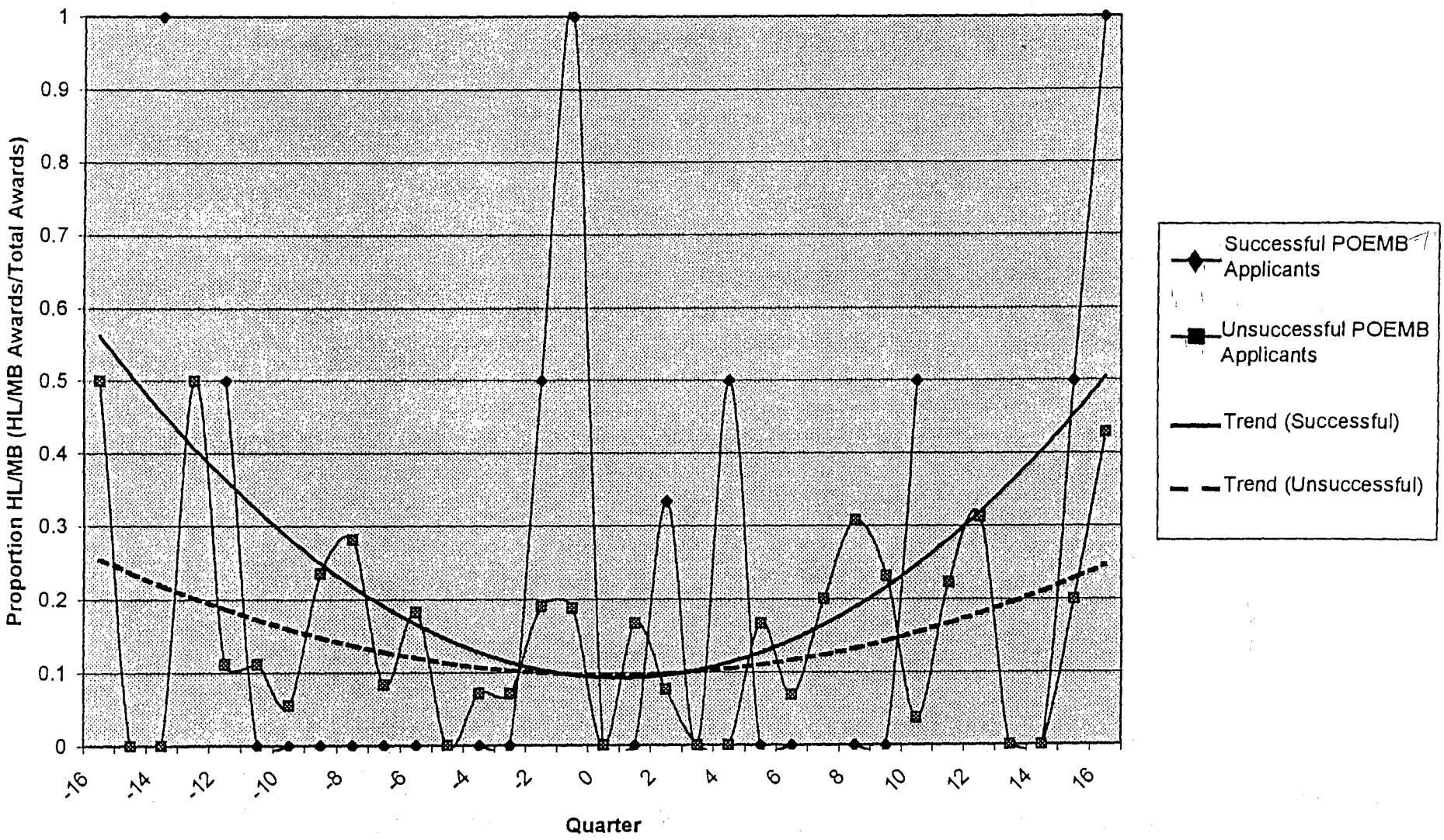


Figure 4.14 Mean Proportion of Total Awards Designated HL/MB, By Quarter



HL/MB-designated awards received by senior investigators in the successful applicant group decreased in the post period relative to the pre period. This difference was not statistically significant, however.

"Other"-designated Awards. Figure 4.15 shows the average number of other-designated awards over time for successful and unsuccessful POEMB applicants, along with the least squares quadratic trend line through each time series. The trends indicate greater numbers of HL awards for unsuccessful POEMB applicants throughout the pre- and post-award periods, consistent with their greater levels of non-HL application activity. The repeated measures analysis showed main effects of program ($p<0.01$), time ($p<0.01$),and time squared ($p<0.01$), but no interaction effect of program by time in the linear or quadratic terms. Of those described thus far, this is the first analysis showing a linear trend over time, in part driven by the very low numbers of "other"-designated awards early in the pre-POEMB observation period (see Figure 4.15). Simple before-after comparisons showed two-fold or greater differences between groups for the aggregate, pre, and post periods (all significant at $p<0.01$). Examination of these data by status showed senior investigators in the successful POEMB applicant group as receiving slightly higher average numbers of other-designated awards for all periods, with a slight increase during the post-period. Potential trainees (SCODE 3) from the unsuccessful POEMB applicant group simply had very few awards during the post period, making a t-test impossible.

Figure 4.16 shows "other"-designated awards as a proportion of total applications over time. Again, caution should be exercised due to high variability (and one extreme value), but the trends appear to suggest a low proportion of "other"-designated awards for both groups early in the pre-award period,followed by a more rapid rise among successful POEMB applicants approaching the post-award period. Both groups' trend lines declined again late in the post-award period, but the successful POEMB applicants maintained the higher rate they had established. Before-after comparisons aggregated across quarters showed a higher average rate for the unsuccessful POEMB applicants in the pre-award period (16 percent to 10 percent), but a higher average rate for the successful applicants in the post-award period (22 percent to 13 percent). Comparisons by status showed a pattern consistent with the previous analysis, with the exception that the proportion for senior investigators in the successful POEMB applicant group is higher than that for the same stratum in the unsuccessful applicant group during the pre period. These differences are not statistically significant, however.

Figure 4.15 Mean Other Awards Per Investigator, By Quarter

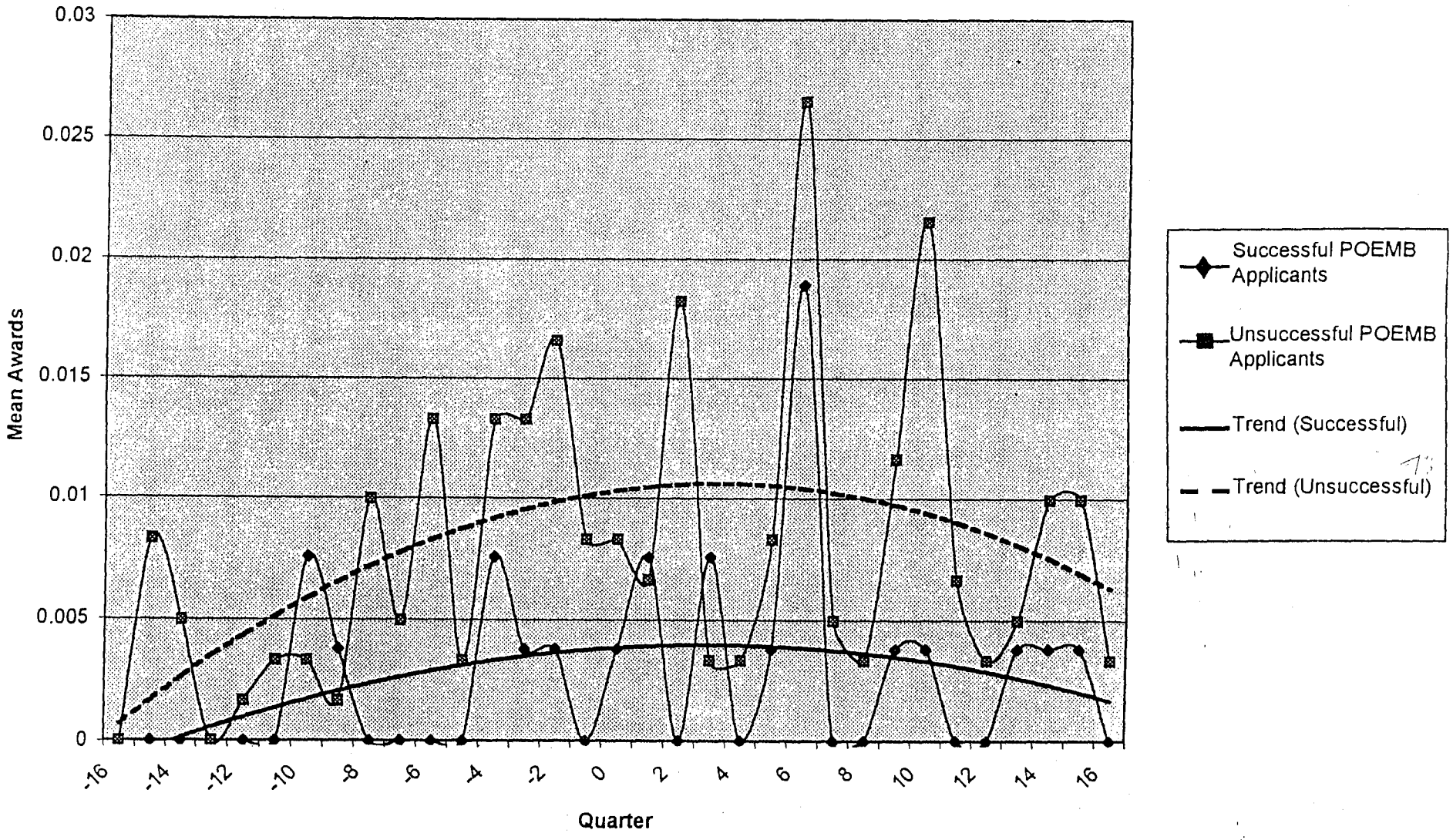


Figure 4.16 Mean Proportion of Applications Yielding Other Awards, By Quarter

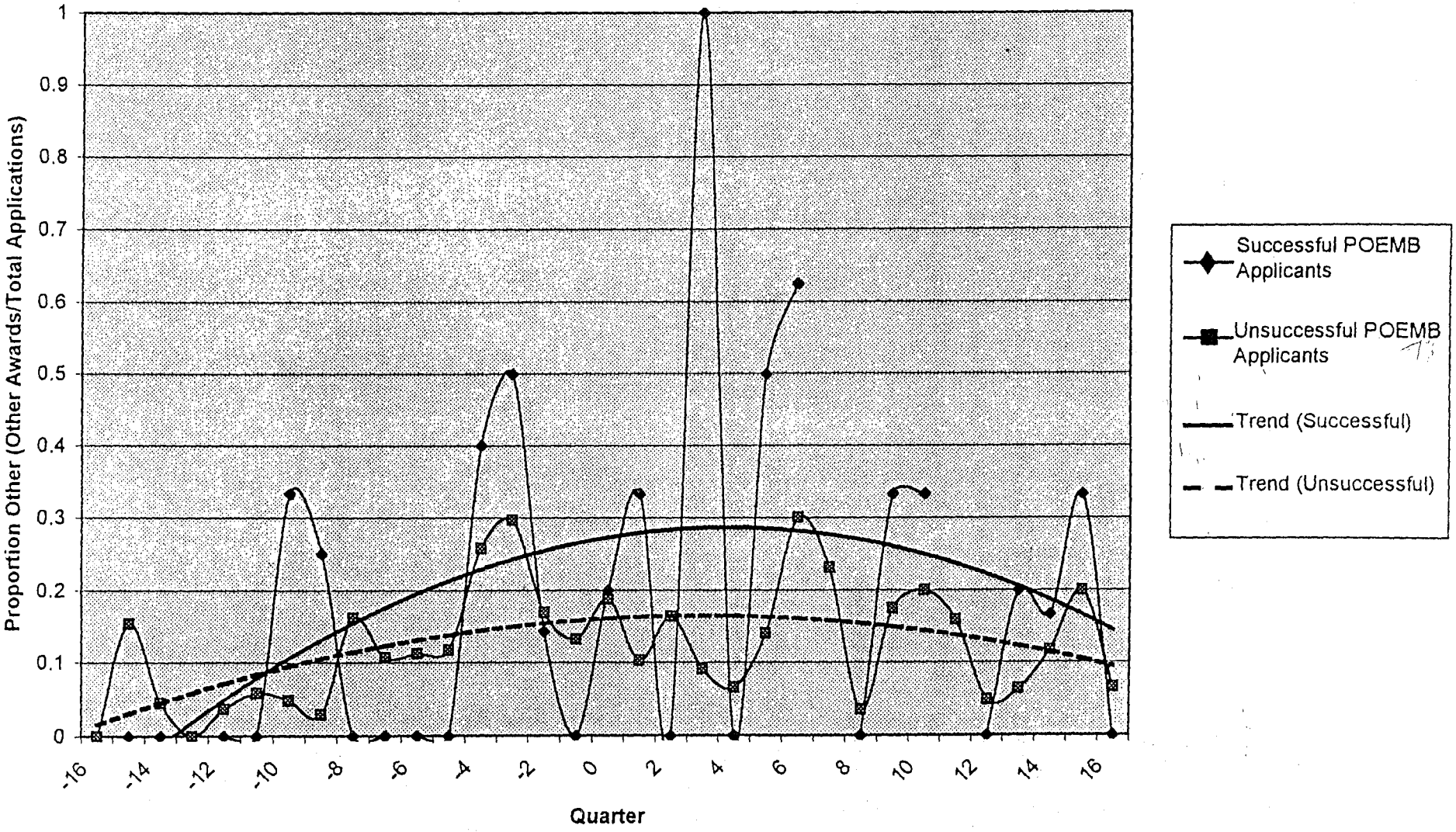


Figure 4.17 shows "other"-designated awards as a proportion of total awards over time. As shown, the trends suggest successful POEMB applicants' average proportions increasing relative to unsuccessful applicants in the post-award period. Consistent with the figure, simple before-after comparisons showed no difference during the pre-award period for investigators receiving one or more awards during that period, while successful POEMB applicants posted higher proportions in the post-award period (40 percent to 32 percent). This difference and the still larger difference in pre-post gain scores among investigators receiving one or more awards in both periods (7 percent to -6 percent for successful and unsuccessful POEMB applicants, respectively) were statistically nonsignificant due to small *n's* in the successful applicant group. Examination of these data by status showed the same patterns in terms of between-POEMB applicant group differences for senior investigators.

4.2.3.2 Analysis of Switching

Tables 4.5 and 4.6 show the complete transition matrices for successful and unsuccessful POEMB applicants, respectively. The transition matrix shows the probability of all possible transitions from pre-POEMB to post-POEMB status. There are eight possible pre and post states:

- None — no NIH awards
- HL only — only HL awards
- MB only — only MB awards
- HL/MB — HL and MB awards
- "Other" only — only "other" awards
- HL/ "other" — HL and "other" awards
- MB/ "other" — MB and "other" awards
- HL/MB/ "other" — HL, MB, and "other" awards

As with grant applications, each pre-POEMB state has a separate set of transition probabilities that sum to 100 percent. In Table 4.5, for example, among investigators who received no grants pre-award, 94 percent received no grants post-award, 1 percent received only "other" grants, 3 percent received only MB grants, 1 percent received HL/MB grants, 1 percent received HL only, and zero percent received HL and "other" grants.

Figure 4.17 Mean Proportion of Total Awards Designated Other, By Quarter

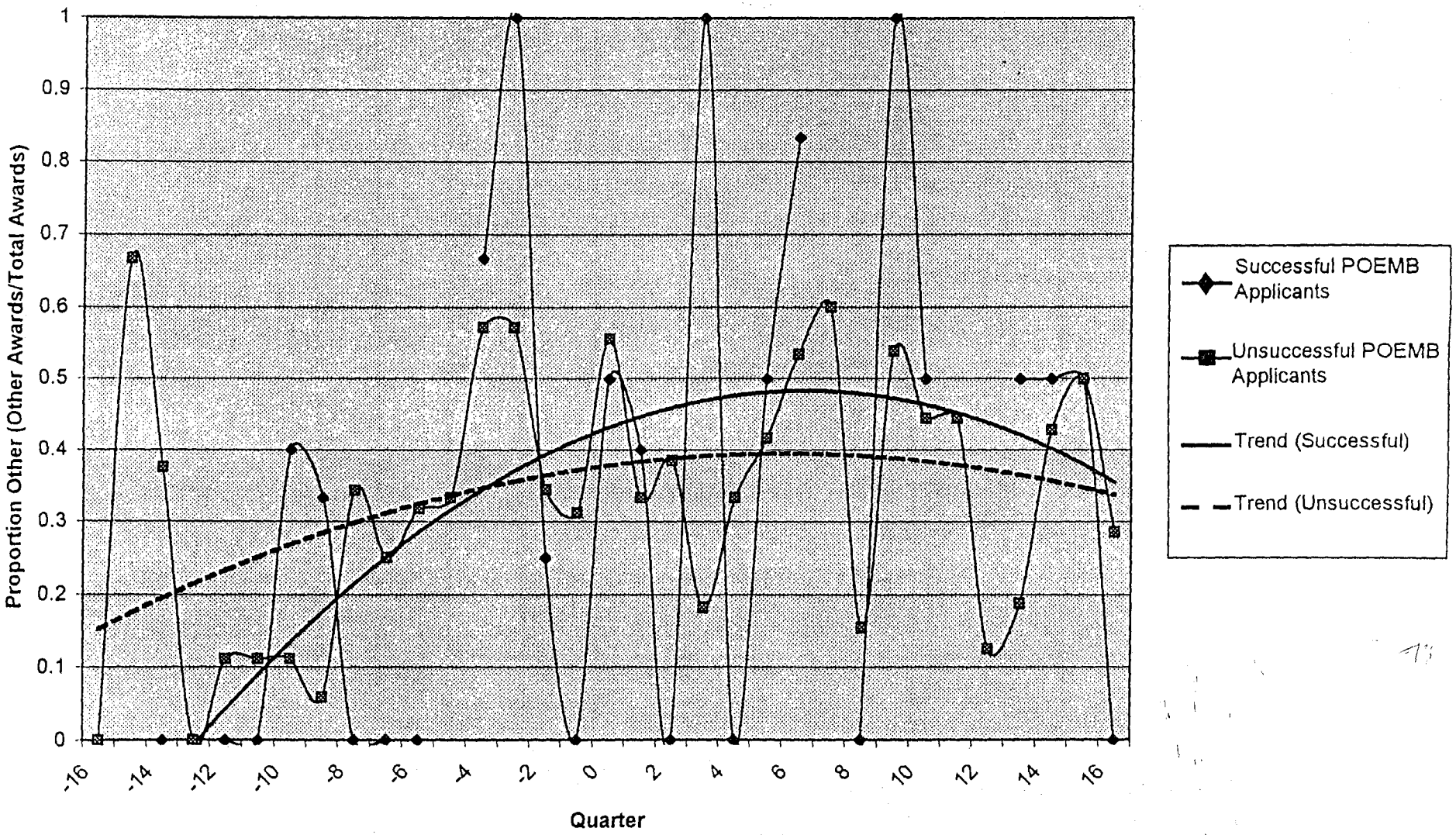


Table 4.5. Pre-POEMB Application to Post-POEMB Application Transition Matrix for NIH/PHS Grant Awards Won by Successful POEMB Applicants

Pre-POEMB Application	Post-POEMB Application Frequency, Row %								
	no awards	"other" awards	MB	HL	HL/MB	MB/ "other"	HL/ "other"	HL/MB/ "other"	Total
no awards	222 94%	3 1%	6 3%	2 1%	2 1%	0 0%	0 0%	0 0%	236
"other" awards	7 70%	1 10%	2 20%	0 0%	0 0%	0 0%	0 0%	0 0%	10
molecular biology (MB)	6 75%	0 0%	1 13%	0 0%	0 0%	1 13%	0 0%	0 0%	8
MB/"other"	1 50%	0 0%	0 0%	0 0%	0 0%	1 50%	0 0%	0 0%	2
HL	1 100%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	1
HL/MB	6 86%	1 14%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	7
HL/MB/ "other"	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0
HL/ "other"	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0
Total %	243 92%	5 2%	9 3%	2 1%	2 1%	2 1%	0 0%	0 0%	264

Table 4.6. Pre-POEMB Application to Post-POEMB Application Transition Matrix for NIH/PHS Awards Won by Unsuccessful POEMB Applicants

Pre-POEMB Application	Post-POEMB Application Frequency, Row %								
	no awards	"other" awards	MB	HL	HL/MB	MB/ "other"	HL/ "other"	HL/MB/ "other"	Total
no awards	345 80%	17 4%	26 6%	23 5%	12 3%	5 1%	2 0.5%	1 0.2%	431
"other" awards	26 58%	6 13%	6 13%	0 0%	4 9%	1 2%	0 0%	2 4%	45
molecular biology (MB)	29 57%	6 12%	5 10%	5 10%	0 0%	5 10%	0 0%	1 2%	51
MB/"other"	5 63%	2 25%	0 0%	0 0%	0 0%	1 13%	0 0%	0 0%	8
HL	14 40%	3	1 3%	9 28%	5 14%	0 0%	2 6%	1 3%	35
HL/MB	13 57%	9%	2 9%	2 9%	3 13%	1 4%	0 0%	0 0%	23
HL/MB/ "other"	2 50%	0 0%	1 25%	1 25%	0 0%	0 0%	0 0%	0 0%	4
HL/ "other"	2 33%	2 33%	0 0%	1 17%	0 0%	0 0%	1 17%	0 0%	6
Total %	436 72%	38 6%	41 7%	41 7%	24 4%	13 2%	5 1%	5 1%	603

The darkly shaded cells represent the 16 transition states that could be categorized as "switching," so designated because each represents a transition from a non-HL to HL state or a non-MB to MB state. To cover all transitions of potential interest, we characterized switching under two definitions: a more restrictive and a less restrictive. The more restrictive or "pure" switch conditions are those in which the investigator transitions from an MB to an HL state or vice versa. These are:

- **HL to MB.** This includes HL only to MB only, HL only to MB/other, HL/other to MB only, HL/other to MB/other.
- **HL to HL/MB.** This includes HL only to HL/MB, HL only to HL/MB/other, HL/other to HL/MB, and HL/other to HL/MB/other.
- **MB to HL.** This includes MB only to HL only, MB only to HL/other, MB/other to HL only, MB/other to HL/other.
- **MB to HL/MB.** This includes MB only to HL/MB, MB only to HL/MB/other, MB/other to HL/MB, MB/other to HL/MB/other.

As shown in Tables 4.5 and 4.6, only a small number of investigators fell into any of the four "pure" switch transitions (13 altogether), and all were in the unsuccessful POEMB applicant group.

The lightly shaded cells in the tables represent the additional less restrictive switch conditions. These were defined as follows:

- **No HL to some HL.** This includes none to HL only, none to HL/MB, none to HL/other, none to HL/MB/other, MB only to HL only, MB only to HL/MB, MB only to HL/other, MB only to HL/MB/other, other to HL only, other to HL/MB, other to HL/other, other to HL/MB/other, MB/other to HL only, MB/other to HL/MB, MB/other to HL/other, MB/other to HL/MB/other.
- **No MB to some MB.** This includes none to MB only, none to HL/MB, none to MB/other, none to HL/MB/other, HL only to MB only, HL only to HL/MB, HL only to MB/other, HL only to HL/MB/other, other to MB only, other to HL/MB, other to MB/other, other to HL/MB/other, HL/other to MB only, HL/other to HL/MB, HL/other to MB/other, HL/other to HL/MB/other.
- **No HL/MB to some HL/MB.** This includes none to HL/MB, none to HL/MB/other, HL only to HL/MB, HL only to HL/MB/other, MB only to HL/MB, MB only to HL/MB/other, other to HL/MB, other to HL/MB/other, HL/other to HL/MB, HL/other to HL/MB/other, MB/other to HL/MB, MB/other to HL/MB/other.

These definitions are less restrictive in that investigators with and without pre-POEMB

application activity are grouped together. As with applications, it is arguable that those who receive no awards in the pre-POEMB period are not truly "switching" when they receive awards in the post-POEMB period; rather they are simply beginning to receive awards in a particular field. They are worth examining, however, particularly since the small number of successful POEMB applicants receiving NIH/PHS awards in the pre-POEMB period (21) greatly limits their potential to demonstrate switching on this indicator.

Of the 867 investigators, 51 met the No HL to some HL condition. Unsuccessful POEMB applicants were more than twice as likely to fall into this subset as were successful applicants (7 percent to 3 percent), and this difference was statistically significant ($p < 0.05$). Among the subset of investigators who began in the No HL condition, (and therefore had the opportunity to switch), the relative frequencies of the two groups showed a similar pattern (8 percent and 3 percent for unsuccessful and successful POEMB applicants, respectively).

More investigators (81) met the No MB to some MB condition. Again, unsuccessful POEMB applicants were over twice as likely to meet this condition (11 percent to 5 percent, $p < 0.01$). Among the subset of investigators who began in the No MB condition, the relative frequencies of the two groups showed a similar pattern (13 percent and 6 percent for unsuccessful and successful POEMB applicants, respectively).

As expected, a smaller number (31) met the No HL/MB to some HL/MB condition. Again, the difference favored unsuccessful POEMB applicants, but was smaller (4 percent to 3 percent) and statistically nonsignificant. Among the subset of investigators who began in the No HL/MB condition, the relative patterns were identical (4 percent to 3 percent).

In sum, there are several switch patterns in which unsuccessful POEMB applicants transitioned at higher rates than did successful POEMB applicants, others in which there was no significant difference, and none in which successful POEMB applicants transitioned at higher rates than unsuccessful applicants. Thus, these data offer no evidence that receiving a POEMB grant increased the probability of investigator switches of the sort hypothesized, at least as measured by the receiving of NIH/PHS grants.

4.3 Awards—American Heart Association (AHA)

This section focuses on an examination of AHA award activity of individuals in the successful and unsuccessful applicant groups. Below we discuss first general AHA awards, followed by a discussion of the AHA Bugher awards.

4.3.1 General AHA Awards

This section is subdivided into (1) methods and (2) findings.

4.3.1.1 Methods

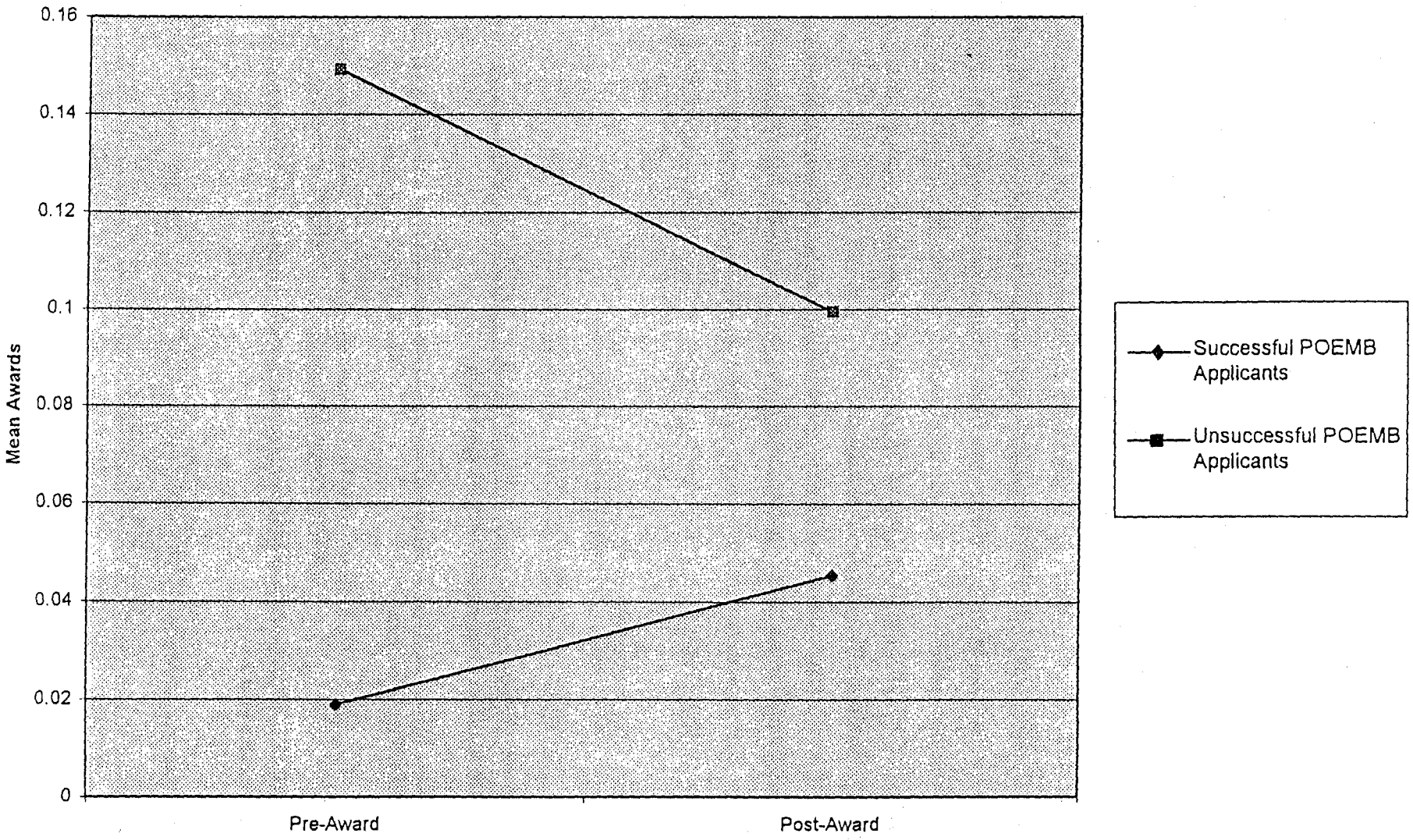
The AHA baseline file was used for the analysis reported on in this section. The analysis of the AHA awards data was carried out using the same methods as described in the previous section of this report with the exception that fewer between-POEMB applicant group comparisons were possible due to a lack of information that would characterize the awards (no descriptors available for this data).

The (null) hypothesis tested in this section is that there are no (statistically significant) differences between the two POEMB applicant groups in terms of the measures used in this analysis (mean differences in total AHA awards, post-POEMB mean differences in total awards versus pre-POEMB mean differences, and mean differences in "switches"). We expect that having received a POEMB grant may have stimulated POEMB researchers to apply for and receive AHA awards — i.e., an increase in the number of AHA awards received by successful POEMB applicants during the post-POEMB application period relative to the unsuccessful POEMB applicants.

4.3.1.2 Findings

Unsuccessful POEMB applicants received significantly more AHA awards over the total period, pre-POEMB period, and post-POEMB period. However, the successful POEMB applicants' mean increased across the two periods while the unsuccessful applicants' mean declined (see Figure 4.18), and this difference between gain scores was statistically significant ($p < 0.05$).

Figure 4.18 Mean Number of AHA Awards Per Investigator, By Quarter



Consistent with their receiving greater numbers of awards, unsuccessful POEMB applicants were significantly more likely to receive one or more AHA awards during the pre-POEMB period (10 percent versus 2 percent, $p<0.01$). This difference persisted but narrowed during the post-POEMB period (7 percent versus 4 percent, $p<0.05$). The proportion of successful and unsuccessful POEMB applicants moving from no awards in the pre-POEMB period to at least one award in the post-POEMB period was the same in both groups (3 percent). However, the proportion moving in the opposite direction differed (1 percent among successful POEMB applicants versus 6 percent among unsuccessful applicants), and this difference was statistically significant ($p<0.01$). That is, the narrowing of the difference between groups in the post-POEMB period appears to be primarily attributable to active investigators in the unsuccessful POEMB applicant group becoming inactive, rather than to inactive investigators in the successful POEMB applicant group becoming active.

Tables 4.7 and 4.8 show the AHA award transition matrices for successful and unsuccessful POEMB applicants, respectively. As with NIH applications and awards, the transition matrix shows the probability of all possible transitions from pre-POEMB to post-POEMB status. In this case, however, the matrix is greatly simplified because there are only two pre-POEMB and post-POEMB states:

- No — no AHA award
- Yes — one or more AHA award

Table 4.7 Pre-POEMB Application to Post-POEMB Application Transition Matrix for AHA Awards Received by Successful POEMB Applicants

Pre-POEMB Application	Post-POEMB Application Frequency, Row %		
	received one or more awards	did not receive any awards	Total
received one or more awards	1 25%	3 75%	4
did not receive any awards	9 3%	251 97%	260
Total Percent	10 4%	254 96%	264

Table 4.8. Pre-POEMB Application to Post-POEMB Application Transition Matrix for AHA Awards Received by Unsuccessful POEMB Applicants

Pre-POEMB Application	Post-POEMB Application Frequency, Row %		
	received one or more awards	did not receive any awards	Total
received one or more awards	23 40%	35 60%	58
did not receive any awards	21 4%	524 96%	545
Total	44 7%	559 93%	603

This time only one transition can be characterized as "switching," hence, the single darkly shaded cell. As shown, successful POEMB applicants with no pre-POEMB AHA awards had virtually the same probability of receiving AHA awards in the post-POEMB period as did their counterparts in the unsuccessful applicant group (3 percent and 4 percent, respectively).

The pre-post period narrowing of the differences between successful and unsuccessful POEMB applicants shown in Figure 4.18 is, on its face, consistent with a hypothesis that receiving a POEMB grant stimulated the application for and the receiving of AHA awards. That the narrowing was primarily due to active investigators in the unsuccessful POEMB applicant group becoming inactive, rather than inactive investigators in the successful applicant group becoming active, muddies the interpretation. These switching data offer no evidence that receiving a POEMB grant increased the probability of receiving an AHA award in the post-POEMB period among investigators who had not received such awards in the pre-POEMB period. It is important to note that the files Battelle received from AHA for analysis extended through 1993, which represents only a portion of the post-POEMB award period.

4.3.2 AHA-Bugher Awards

In 1985, the Henrietta B. and Frederick H. Bugher Foundation and the American Heart Association (AHA) formed a partnership to establish a group of Centers for Molecular Biology in the Cardiovascular System. The goal of this partnership was to initiate an integrated, institution-based

research program to train promising individuals with cardiovascular medicine backgrounds in molecular biology research techniques and to promote the development of a focus of research work in the application of molecular biology to the study of components of the cardiovascular system (Morgan and Paul, 1995) — a goal similar to that of NHLBI's POEMB. This similarity warranted the consideration of participation in the AHA-Bugher Award as a confounder in the evaluation of the effect of POEMB on the grant and publication activity of the applicants.

4.3.2.1 Methods

Data on the fellows who participated in the AHA-Bugher Foundation Centers for Molecular Biology in the Cardiovascular System were provided by the AHA's Division of Research Administration. This information included the first name, last name, and middle initial of Bugher fellows as well as whether he/she was funded by AHA in the fiscal year(s) in which they participated. In some cases, an individual might be called a Bugher fellow but not actually receive funds from AHA in that year.

The AHA-Bugher fellows data were manually cross-matched with the complete list of names from the **PMBPERS** file in order to identify individuals who were both POEMB applicants (whether successful or unsuccessful) and AHA-Bugher fellows. The PCODEs of these individuals are listed in Table 4.9.

4.3.2.3 Findings

Table 4.9 also presents information on the overlap of successsful and unsuccessful POEMB applicants and AHA-Bugher fellows. Only one investigator from the successful applicant group received funding from AHA-Bugher (0.3 percent overlap). However, for this individual, there was no temporal overlap in funding from the two organizations. There was a greater degree of overlap among the unsuccessful POEMB applicant group (1.5 percent overlap) with nine individuals having participated as AHA-Bugher fellows. It should be noted, however, that three of these individuals were listed as having been AHA-Bugher fellows but *not* as having received AHA funds. Excluding these individuals from consideration yields a 1 percent overlap for the unsuccessful applicant group.

Table 4.9. Overlap in Investigators between AHA-Bugher Awards and POEMBs

Successful POEMB Applicant Group	AHA-Bugher Institution	PCOD E	Percent Overlap	Timing of AHA-Bugher	POEMB Application Date
Status 4	Baylor	1.46	1/246=0.3% by status overlap of 25%	1986-87	1988
Unsuccessful POEMB Applicant Group					
Status 2	Baylor	11.05	9/603=1.5% by status overlap of 1.6%	1986-88	1988 and 1989
	Baylor	11.15		1986-88	1988 and 1989
	Children's Hospital	12.06		1987-89	1989
	Children's Hospital	12.12*		1987-88*	1989
	Children's Hospital	12.18*		1987-88*	1989
	U. Texas SW Medical Center	14.21*		1987-88*	1989
Status 3	Children's Hospital	13.48	by status overlap of 1.4%	1989-90	1988
	Children's Hospital	13.49		1988-89	1988
	Children's Hospital	13.63		1989-91	1988

* Individuals listed by AHA-Bugher as fellows but who did not actually receive funds from AHA.

The overlap occurred prior to the date of POEMB application for all but three of the unsuccessful POEMB applicants who were also listed as AHA-Bugher fellows. For those three potential trainees, it is possible that the experience of having applied for POEMB may have somehow increased their chances of receiving AHA-Bugher funding. This possibility cannot be assessed here, however, as the names of the unsuccessful AHA-Bugher applicants were not made available for analysis. (This may not have been possible to determine at all since the *n* is so small.) The grant application activity of the individual from the successful POEMB applicant group listed in Table 4.9 was determined for the total period (pre- and post-POEMB). This researcher submitted no grant applications during this period. The nine unsuccessful POEMB applicants who were listed as Bugher fellows submitted a total of 14 grant applications during the total period (17 quarters pre- and 16 post-POEMB application). Overall, the absence of a temporal overlap for the one overlapping successful POEMB applicant and the lack of grant application activity for this individual rules out the possibility that the AHA-Bugher Award may have acted as a confounder in the analysis described in the previous sections of this chapter.

4.4 Missionary Impact of POEMB on Those Investigators and Trainees Who Left the POEMB

This section focuses on an examination of the missionary impact of POEMB participation on those who left the program midway through the grant period. The section is subdivided into (1) methods and (2) findings.

4.4.1 Methods

In order to determine whether the investigators and trainees who left the POEMB continued to pursue applications of molecular biology in the fields of cardiovascular and pulmonary disease, their subsequent grant application activity was described. Data utilized in the previous sections to assess grant application and award activity were used in this examination (**CGAF** baseline file). The variables _1 through _7 from the **PMBPERS** file were used to identify those investigators and trainees who left the program and determine the timing of their departure. These seven variables represent possible data

sources for information contained in each record of **PMBPERS**: _1 (= the original application), _2 (= the first progress report), _3 (= the second progress report), etc. The fields that correspond to the document(s) in which the individual is cited are tagged with an "X" as a means of tracking POEMB investigators and trainees throughout the POEMB period. University of California at San Fransisco (UCSF) began their POEMB in 1989. For the purposes of this analysis (because 1994 was the last year that we received information on these individuals), _6 was the last year that the UCSF individuals were tracked.

If the POEMB had an impact on the research careers (in terms of direction and focus) of those who departed the program before the end of the first funding period, we would expect to observe these individuals (1) submitting HL-designated grant applications and/or (2) receiving awards that are associated with HL and MB-designated **CRISP** terms.

4.4.2 Findings

Table 4.10 presents information on the grant activity of those researchers who left POEMB before the end of the program.[7] A total of 126 researchers left the POEMB before the end of the program, most of whom were trainees (SCODE 4). A total of 24 applications were submitted by these individuals during the period after their departure from the POEMB. Four of the 24 applications (17%) were classified as heart/lung-related using their associated ICD codes. Fourteen of the 24 applications (58%) were awarded, but only five (36%) were associated with HL-designated **CRISP** descriptors. The remainder of the awards were associated with "other"-designated descriptors.

In sum, the data are too sparse to provide clear evidence that the individuals who left the POEMB before the end of the program continued to pursue applications of molecular biology in cardiovascular and/or pulmonary disease research. It is important to note, however, that trainee-level (SCODE 3 and 4) are responsible for all of the post-POEMB heart/lung research activity, at least as measured by ICD code and **CRISP** terms. This pattern is consistent with the hypothesis that the research activity of POEMB trainees may have been stimulated by the program, leading them to pursue funding for new research projects focused on cardiovascular and/or pulmonary research problems even

---

[7] In the interest of preserving the confidentiality of the POEMB investigators, we present the publication data in aggregate.

after their departure from the POEMB.

Table 4.10 Post-departure Grant Activity of Researchers Who Left POEMB

Number of Key Investigators and Trainees Who Left POEMB	Number of Grant Applications Submitted Following Departure	Number of HL-designated (ICD) Grants Submitted Following Departure	Number of Awards Received Following Departure and Associated CRISP Terms
Status 1 = 1	0	0	0
Status 2 = 4	4	0	0
Status 3 = 32	6	0	HL only = 1 "other" = 8
Status 4 = 89	14	4	HL only = 1 "other" = 1
Total = 126	Total = 24	Total = 4	

4.5 Summary of Findings

In this chapter, we focused on the grant application and award activity of the individuals in the two POEMB applicant groups as measures of their research activity. Using the **CGAF/CRISP** baseline file, we were able to make comparisons between the two POEMB applicant groups in terms of: (1) total grant application submissions, (2) the research focus of grant applications as measured by ICD — non-heart/lung (nonHL) versus heart/lung (HL), (3) total awards received, and (4) the research focus of awarded grants as measured by **CRISP** descriptor categories.

4.5.1 NIH/PHS Applications

Below we summarize the findings on the NIH/PHS grant applications activity of individuals in the successful and unsuccessful POEMB applicant groups for (1) longitudinal analysis and (2) switching analysis.

4.5.1.1 Longitudinal Analysis

An examination of the raw frequencies of total grant applications submitted by the individuals in the two POEMB applicant groups showed that the most apparent pre- to post-POEMB differences in grant application activity were:

- The decrease in total number of applications submitted by senior investigators in the successful POEMB applicant group, and
- The more than two-fold increase in total submissions (10 to 25) exhibited by the trainee-level researchers from the successful POEMB applicant group.

Because a similar decrease in the total number of applications was not observed among senior investigators in the unsuccessful POEMB applicant group, the former difference is consistent with the hypothesis that POEMB shields the investigators from having to apply for multiple grants to maintain funding streams. At first glance, the latter pre-post difference might appear to represent "maturation," i.e., the fact that junior investigators will become more active in submitting grant applications as their research careers develop. Had the effect been simply due to maturation, a similar increase would be observed among potential trainees in the unsuccessful POEMB applicant group; however, it was not. Thus, this observed effect is consistent with the hypothesis that receiving the POEMB afforded trainees the opportunity and provided the stimulus to pursue more grants independently than would normally be expected of trainee-level researchers.

In terms of the average number of grant applications submitted per investigator, unsuccessful POEMB applicants (all statuses) submitted more than four times as many total grant applications as did successful POEMB applicants (all statuses), and almost four times as many NHLBI applications, during the total period, the pre-POEMB period, and the post-POEMB period. However, examination of these same data by investigator status (senior investigators versus trainees) eliminated the four-fold difference between the average total grant submissions of the senior investigators from the two applicant groups for all three periods. This is because the four-fold difference in the pooled data was largely driven by the far higher percentage of trainees in the successful applicant group (87 percent versus 37 percent), as described in the methods section. Still the difference between the average number of total grants submitted by senior investigators in the two applicant groups remained statistically significant ($p<0.01$)

with successful senior investigators submitting fewer grant applications overall and decreasing relative to the unsuccessful senior investigators for the post-POEMB period. For the pre-POEMB period, a comparison of senior investigators from both POEMB applicant groups showed no difference in the average number of grants submitted. This is consistent with the hypothesis that participation in the POEMB has freed the senior investigators in the successful applicant group from having to apply for multiple grants to maintain streams of funding.

The longitudinal analysis of the ratio of HL-designated applications to total applications yielded trends for the two POEMB applicant groups over the entire 33-quarter period.

- The unsuccessful POEMB applicant group's trend is practically flat, while the successful applicant group's ratio starts lower and ends higher than that of the unsuccessful applicant group — a classic group-by-time interaction. This indicates that, controlling for total application activity, the successful POEMB applicants' HL ratio increased relative to the ratio for the unsuccessful applicants.

- Correspondingly, simple before-after comparisons with data aggregated across quarters showed the unsuccessful POEMB applicant group with a higher average ratio of HL to total applications in the pre-award period (41 percent for unsuccessful to 32 percent for successful), while the successful applicant group had a higher average ratio of HL to total applications in the post-award period (49 percent for successful to 33 percent for unsuccessful).

- When examined by status, however, a different pattern was exhibited. Once stratified, there was no difference between the ratios of the senior investigators in the two POEMB applicant groups, while the POEMB trainees increased pre to post POEMB and the potential trainees of the unsuccessful group decreased pre to post POEMB. This difference was not statistically significant, however.

Although not statistically significant, the difference between the trainees in the two applicant groups is consistent with the hypothesis that receiving the POEMB stimulated trainees to pursue more grants, particularly those designated as HL.

4.5.1.2 Analysis of "Switching"

Analysis of "switching" patterns, in which we examined the effect of POEMB on the focus of the research of the individuals in the two applicant groups as evidenced by the submission of grant applications, yielded interesting results. For this analysis, the ICD code associated with the grant

applications was used to determine whether there was a shift or switch from pre-POEMB states in which heart/lung designated applications were *not* being submitted to post-POEMB states in which heart/lung designated applications *were* being submitted. A comparison of these two applicant groups suggests that the POEMB may indeed contribute to such a shift. Specifically, the ICD switching analysis showed:

- Considering all statuses together, successful POEMB applicants submitting only non-HL applications in the pre-award period had a higher probability of submitting only HL applications in the post-award period. Unsuccessful applicants had a higher probability of transitioning from the non-HL only state to the mixed state. These differences are statistically significant ($p < 0.01$).

- Aggregating across all four switching categories, unsuccessful POEMB applicants were more likely to switch, and this difference was statistically significant ($p < 0.05$).

- Examining the subset of investigators who were not submitting HL-designated applications in the pre-POEMB period separately showed that individuals from the unsuccessful POEMB applicant group were two times as likely as those from the successful applicant group to switch to applying for HL-designated applications in the post-POEMB period ($p < 0.01$).

- Examining only the subset of researchers who were active in applying for grants during the pre-POEMB period showed that successful POEMB applicants were again slightly more likely to make this switch from a state in which HL-designated grants were not applied for to a state in which they were applied for than was the case for unsuccessful POEMB applicants, although the difference was not statistically significant.

- Examining the subset of individuals who were active in the pre-POEMB period *and* capable of switching (this excludes those who were inactive or already submitting HL applications in the pre-POEMB period) showed that, again, successful POEMB applicants submitting only nonHL-designated applications in the pre-POEMB period had a higher probability of submitting only HL-designated applications in the post-POEMB period. This difference was not statistically significant, however.

In sum, there were several switch patterns in which unsuccessful POEMB applicants transitioned at significantly higher rates than did successful applicants, others in which there was no significant difference, and one in which successful POEMB applicants transitioned at significantly higher rates than did applicants in the unsuccessful applicant group. That the successful POEMB applicants made the switch from non-HL only to HL only at significantly higher rates than the unsuccessful applicants is consistent with the hypothesis that receiving the POEMB may have led to a shift or switch in focus from non-HL research to HL research as measured by grant application activity and ICD. The findings from

the ICD switching analysis provided no other clear evidence that receiving a POEMB grant increased the probability of investigator switches of the type hypothesized, at least as measured by the application for NIH/PHS grants. It is important to note that the observed difference was only evident when a less restrictive definition of switching categories was employed.

4.5.2 NIH/PHS Awards

The receiving of NIH/PHS grant awards was examined for the successful and unsuccessful applicant groups, with analyses based on longitudinal patterns and on comparisons between these two groups in terms of the total awards received over time (longitudinal) and the research focus of receiving grants as measured by **CRISP** descriptor categories ("switching").

4.5.2.1 Longitudinal Analysis

The longitudinal analysis of NIH/PHS awards focused on detecting specific types of trends in data related to achievement of the objectives of the POEMB program.

- **Total Awards.** The longitudinal trends are similar for each type of **CRISP** term-designated awards (e.g., HL, MB, HL/MB, "other") in terms of the average number of total awards received over time: the unsuccessful POEMB applicants received a greater number of awards throughout the pre- and post-POEMB periods, consistent with their greater levels of application activity. However, this difference in the pooled data was driven by the far higher percentage of trainees in the successful applicant group. The difference disappeared when the data were examined separately by status code, but still supplied no evidence consistent with an effect of POEMB on total awards at the investigator level.

- **Success Rate.** The trends of average success rate over time indicate that successful POEMB applicants tended to register higher success rates in the period following POEMB application date than did their counterparts in the unsuccessful POEMB applicant group. Analysis by status showed that it was the trainees from the successful applicant group who were receiving more awards during this period than were the potential trainees from the unsuccessful applicant group. This may be interpreted in part as a combined effect of maturation (of trainees) and shielding (of senior investigators) from pressure to apply for grants.

- **Heart and Lung-designated Awards.** All analyses showed that unsuccessful POEMB

applicants received more HL-designated awards than did successful applicants.

- **Molecular Biology-designated Awards.** Before-after comparison of the average number of MB-designated awards by status showed that senior investigators from the successful POEMB applicant group received a greater average number of MB-designated awards, with a slight decrease from the pre to the post period, although this difference was not statistically significant. In terms of MB-designated awards as a proportion of total applications over time, successful POEMB applicants had a higher average proportion for both periods, though again the differences were not significant. Furthermore, successful POEMB applicants had a higher proportion of molecular biology-designated awards in their "portfolios" (i.e., MB-designated awards as a proportion of total awards) for the pre-POEMB period, but this difference disappeared in the post-POEMB period.

- **Heart and Lung/Molecular Biology-designated Awards.** Successful applicants posted higher proportions of HL/MB-designated awards in the post-POEMB period compared to the unsuccessful applicants, both in terms of proportion of total applications and total awards over time. However, neither of the before-after comparisons for these two proportions were statistically significant. The pattern of the trend lines for the proportion of total applications is perhaps the clearest yet presented of an effect that is consistent with the hypothesis that receiving a POEMB grant affected investigator behavior. The two trend lines were flat and virtually identical throughout the pre-award period, whereas only the successful POEMB applicants' line increased in the post-award period. Examination of these data by status eliminated the different rates for all periods for senior investigators. This suggests that it is the awards received by the trainee-level researchers in the successful POEMB applicant group that are affecting the difference in the averages in the pooled pre-post comparison discussed above. Such a pattern is consistent with the hypothesis that the trainees are maturing, striking out on their own, and integrating molecular biology and cardiovascular/pulmonary research in their work.

- **"Other"-designated Awards.** Examination of the data by status showed senior investigators in the successful POEMB applicant group as receiving slightly higher average numbers of "other"-designated awards for all periods, with a slight increase during the post-period. In terms of "other"-designated awards as a proportion of both total applications over time and total awards over time, the trends appear to suggest that successful POEMB applicants' average proportions are increasing relative to those of unsuccessful applicants in the post-award period. Before-after comparisons aggregated across quarters for both proportions showed a higher average rate for the successful applicants in the post-award period. In neither case were these differences statistically significant.

4.5.2.2 Analysis of "Switching"

In this analysis, we focused on switching patterns using two types of **CRISP** descriptor-

designated transition categories: a more restrictive and a less restrictive. The more restrictive or "pure" switch conditions are those in which the investigator transitions from an MB to an HL state or vice versa, while the less restrictive switch conditions are those in which investigators with and without pre-POEMB grant application activity are grouped together. A comparison of these two applicant groups showed the following:

- Only a small number of investigators fell into any of the "pure" switch transitions (HL to MB, HL to HL/MB, MB to HL, and MB to HL/MB), and all were in the unsuccessful POEMB applicant group.
- In terms of the less restrictive definitions of switching, unsuccessful POEMB applicants were more than twice as likely to switch from no HL to some HL awards as were successful applicants, a difference that was statistically significant ($p < 0.05$).
- In terms of the less restrictive definitions of switching, unsuccessful POEMB applicants were more than twice as likely to switch from no MB to some MB awards ($p < 0.01$).
- In terms of the less restrictive definitions of switching, unsuccessful POEMB applicants were more likely to switch from no HL/MB to some HL/MB awards, but the difference was smaller than for other switching conditions and was thus statistically nonsignificant.

In sum, there are several switch patterns in which unsuccessful POEMB applicants transitioned at higher rates than did successful POEMB applicants, others in which there was no significant difference, and none in which successful POEMB applicants transitioned at higher rates than unsuccessful applicants. Thus, these data offer no evidence that receiving a POEMB grant increased the probability of investigator switches of the sort hypothesized, at least as measured by the receiving of NIH/PHS grants.

4.5.3 AHA Awards

The receiving of AHA awards was examined for successful and unsuccessful applicant groups. Regarding the AHA awards, both general and Bugher, the analyses showed that:

- The mean number of AHA awards per investigator for the successful POEMB applicant group increased across the two periods while the unsuccessful applicants' mean declined, a statistically significant difference ($p < 0.05$).

- Unsuccessful POEMB applicants were significantly more likely to receive one or more AHA awards during the pre-POEMB period ($p < 0.01$), a difference that narrowed but persisted at a significant level ($p < 0.05$) during the post-POEMB period.

- The narrowing of the difference between the two applicant groups in the post-POEMB period appears to be primarily attributable to active investigators in the unsuccessful POEMB applicant group becoming inactive, rather than to inactive investigators in the successful POEMB applicant group becoming active.

- In terms of "switching," the proportion of successful and unsuccessful POEMB applicants moving from no AHA awards in the pre-POEMB period to at least one AHA award in the post-POEMB period was the same in both groups. The switching data offer no evidence that receiving a POEMB grant increased the probability of receiving an AHA award in the post-POEMB period among investigators who had not received such awards in the pre-POEMB period.

- Participation in the AHA Bugher Award did not prove to be a confounder in the analysis of the grant and award activity of the two POEMB applicant groups.

4.5.4 Missionary Impact

Assessment of the "missionary" impact of those researchers affiliated with a POEMB program at one time but who left before the conclusion of funding did not yield conclusive results. The data are too sparse to assess whether the individuals who left the POEMB before the end of the program continued to pursue applications of molecular biology in cardiovascular and/or pulmonary disease research as measured in terms of grant application and award activity.

4.6 Factors Limiting Analysis

The POEMB's extended award period of seven years was adopted to enable POEMB researchers to pursue new and more innovative developments in molecular biology that could be applied to complex experimental models and systems — a process that would require additional time to develop and use. Given the nature of this type of research, it may be too soon after the start of the POEMB to detect the

kinds of effects hypothesized in the beginning of this chapter. Furthermore, the limited range of the **CGAF/CRISP** baseline file (up to 1993) did not allow us to examine the successful and unsuccessful POEMB applicants' grant and award activity for the entire period of the first POEMBs. Had the data file been extended to 1996, we might have been better able to detect any changes in grant and award activity that might be attributable to the effect of the POEMB.

[**Prompt** interviewee to address each of the two primary objectives.]

- *To foster utilization of molecular biology approaches in important research areas (cardiovascular and pulmonary) within the mission of the NHLBI (technology/research development and application).*
- *To provide opportunities for investigators who have the potential for independent research careers to become skilled in the experimental strategies and techniques of molecular biology and their application to research relevant to the mission of NHLBI (training, dissemination and fostering the NHLBI mission).*

If Yes, then

a. How has your component contributed to this success?

b. Can you give an example of each objective within your component?

c. Was there an emphasis on one objective more than the other, or were the component efforts and resources equally distributed?

6. In what ways did the program not succeed in achieving the primary objectives?

Probe:

a. Which of the objectives were more problematic?

b. Why?

c. What could have been done differently?

e.g., administratively, logistically, personnel, research focus, *etc.*

7. Has your component been successful in achieving it's research goals?

a. Explain.

[Prompt interviewee to address the specifics of his/her component gleaned from the applications and progress reports. Goals will be unique to each component. The companion document(s) should be tailored for each interview.]

Section IV. POEMB Network

1. Did you have opportunities for interaction, collaboration, and the exchange of information with investigators of the other two programs?

a. Explain and give examples.

e.g., personal contacts, Internet communication, conferences, collaboration in publishing, co-presentations at meetings, technology/personnel exchanges, cross-training of personnel, *etc.*

2. If so, was this beneficial to your component's research?

   e.g., increased technology access, other sources of qualified trainees or investigators, sharing of data and resources, resolution of common technical or data analysis problems, synergy in approaches to research, *etc.*

3. If not, would you have found this interaction helpful?

Section V. Recommendations

1. Do you have any **recommendations** for NHLBI regarding the future of POEMB and the pursuit of the primary objectives?

POEMB Primary Objectives:

- *To foster utilization of molecular biology approaches in important research areas (cardiovascular and pulmonary) within the mission of the NHLBI (technology/research development and application).*
- *To provide opportunities for investigators who have the potential for independent research careers to become skilled in the experimental strategies and techniques of molecular biology and their application to research relevant to the mission of NHLBI (training, dissemination and fostering the NHLBI mission).*

6A.3 Interview Guide
Non-POEMB Principal Investigators

Section I. Value of POEMB--Advantages and Disadvantages

1. **Why** did you respond to the initial POEMB RFA?

   Probe:
   a. What specific **aspects** of the POEMB concept prompted you to organize an application and why?

   [Prompt interviewee to address the five key features of POEMB]
   - *The extended seven year award period.*
   - *Cross-disciplinary recruitment and collaboration.*
   - *Opportunity to enhance facilities and resources.*
   - *Broad degree of research freedom--ability to implement innovative approaches.*
   - *Support for new investigators.*

   b. Which of these **aspects** would you consider the most important in prompting you to submit a POEMB application and why?

2. What did you see as the **disadvantages** of the POEMB approach?

Section II. Knowledge of the Programs of Excellence in Molecular Biology

1. What do you think are the most important accomplishments in the application of molecular biology in the areas of cardiovascular and pulmonary diseases in the last several years?

   a. How did you learn about these accomplishments?
   i.e., publications, conference papers

   b. Why do you think they are important?

2. Are you **familiar with** the any of three programs that were granted the POEMB award, and if so, which ones?

   If interviewee not familiar with any of the Programs, then explain which institutions hold the grants.

**Ask Quest. 3 only if interviewee is aware of the Programs:*
3. Are you aware of any **important accomplishments** on the part of the POEMB grantees in applying molecular biology in the areas of cardiovascular and/or pulmonary diseases?

a. If so, what are those accomplishments and how did you learn about them?

*i.e., through **publications or presentations at conferences.***

b. In what ways have these accomplishments contributed to the application of the techniques of molecular biology in the areas of cardiovascular and/or pulmonary diseases?

Probe:
c. Have these programs resulted in the establishment of new research endeavors and/or investigators applying the techniques of molecular biology to the field of cardio-pulmonary disease?

d. In what ways have these accomplishments contributed to the understanding and treatment of cardiovascular and/or pulmonary diseases?

Section III. Continued Interest in Techniques of Molecular Biology

1. Have you or any of your colleagues **maintained an interest** in the application of the techniques of molecular biology in the areas of cardiovascular and/or pulmonary diseases?

If so, ask interviewee to please explain **how they have maintained interest.*

Probes:
a. Was alternative funding sought/granted; and how does it compare with that of the POEMB program?

b. Have you conducted research related to the application of molecular biology techniques in the areas of cardiovascular and/or pulmonary disease?

c. Have you pursued research originally proposed in your POEMB application?

**If interviewee is still working in this area, then proceed to question 2.*

2. Have the significant accomplishments in this area that we discussed earlier **influenced** your own research and ideas?

a. If so, how, in what ways?

**If interviewee was previously aware of any of the three Programs, then proceed to questions 3 and 4.*

3. Do you collaborate or network with individuals in the current POEMB programs?

4. Are the POEMB programs a source of new technology, approaches to cardio-pulmonary research or trained personnel?

Appendix 6B

Code Book and Coding Strategy

Appendix 6B: Code Book and Coding Strategy

The General Code Framework

The framework consists of codes tied to interview questions that are directly related to the two main types of information collected for the study: information on advantages and disadvantages to the POEMB approach (A) and on POEMB accomplishments (B). These are outlined below by the respondent group(s) to which they pertain: POEMB principal investigators, POEMB associate investigators, and/or non-POEMB principal investigators.

A. General data on advantages and disadvantages needed **across all three respondent groups**:

Obtain respondents' perspective as to the advantages and disadvantages of the POEMB approach, with particular emphasis on the aspects of the program that led them to organize an application or to participate in its preparation.

B. Data on accomplishments in two main categories:

1. Data from **POEMB principal investigators** and **POEMB associate investigators**:

Solicit from respondents their views on the most significant accomplishments of their particular grant and on whether and how the POEMB concept contributed to these accomplishments.

2. Data from **non-POEMB principal investigators**:

Ask whether respondents are aware of any important accomplishments of the programs that received the award and whether they or any of their colleagues maintained an interest in the application of the techniques of molecular biology in the areas of cardiovascular and/or pulmonary diseases.

General Codes and the Linked Objectives

Evaluation Objectives and Code Objectives		
A. Obtain investigators' perspectives as to the ***value and desirability*** and the ***advantages*** and ***disadvantages*** of the POEMB approach, with particular emphasis on *the aspects of the program that led them to organize an application or to participate in its preparation.*		
Primary Codes	**Secondary Codes**	**Tertiary Codes**
VALUE Interviewee's perspective on the **value, desirability, and advantages** of the POEMB approach. [Sect. II, #'s 1, 1a, 1b] [Unsucc. Guide: Sect. I, #'s 1, 1a, 1b]	**ASPECT** Discussion of the **aspects** of POEMB that led the investigator to organize an application or to participate in its preparation. Special attention paid to the five key features of the POEMB. [Sect. II, #'s 1a,1b]	**SEVEN:** indicates the extended seven-year award period. **RECRUIT:** indicates the opportunity for cross-disciplinary recruitment and collaboration. **ENHANCE:** indicates the opportunity to enhance facilities and resources. **FREEDOM:** indicates the broad degree of research freedom—ability to implement innovative approaches. **SUPPORT:** indicates support for new investigators. **ALL:** indicates all five of the key features were attractive and important. **NONE:** indicates that none of the key features were attractive and important.
DISADVANT Interviewee's perspective on the **disadvantages** of the POEMB approach, at the time of application and currently. [Sect. II, #2][Unsucc. Sect. I, #2]		

B1. Solicit from investigators on successful POEMB applications their views on the most ***significant accomplishments*** of their particular grant and on whether and how the POEMB concept ***contributed*** to them.		
Primary Codes	**Secondary Codes**	**Tertiary Codes**
ACCOMPLISH Interviewee's discussion of the most **significant accomplishments** of his/her particular Program. This is only to be used for interviewees associated with a Program (i.e., POEMB principal investigators or POEMB associate investigators). [Sect. III, #1]	**OBJECTIVES** Indication of whether and how the program and/or component has been successful in achieving the POEMB primary objectives. [Sect. III, #'s 5 and 6]	**FOSTER:** discussion of whether or not program and/or component has succeeded in fostering the utilization of molecular biology approaches in important research areas within the mission of the NHLBI. [one of the primary objectives]
		OPPORTUN: discussion of whether or not program and/or component has succeeded in providing opportunities for investigators to become skilled in the experimental strategies and techniques of molecular biology and their application to research relevant to the mission of the NHLBI. [one of the primary objectives]
	OWNSUCCESS Indication of whether and how the program and/or component has been successful in achieving its specific goals and objectives. [Sect. III, #'s 7 and 8]	

B1. Solicit from investigators on successful POEMB applications their views on the most ***significant accomplishments*** of their particular grant and on whether and how the POEMB concept ***contributed*** to them.

Primary Codes	Secondary Codes	Tertiary Codes
CONTRIBUTE Interviewee's discussion of whether and how the POEMB concept **contributed** to their Program's significant accomplishments. This can be used with all interviews in order to account for the non-POEMB investigators' assessments of the contribution of POEMB. [Sect. III, #2]	**SEVENCON:** indicates the extended seven-year award period.	
	RECRUITCON: indicates the opportunity for cross-disciplinary recruitment and collaboration.	
	ENHANCECON: indicates the opportunity to enhance facilities and resources.	
	FREEDOMCON: indicates the broad degree of research freedom—ability to implement innovative approaches.	
	SUPPORTCON: indicates support for new investigators.	
	ALLCON: indicates all five of the key features were attractive and important.	
	NONECON: indicates that none of the key features were attractive and important.	
	OUTSIDE: indication of whether or not the accomplishment(s) could have been achieved outside of or without the POEMB. [Sect. III, #3]	
	HINDER: indication of whether any features or aspects of the POEMB concept/structure were a hindrance or obstruction to the program's or component's research and/or goals. [Sect. III, #4]	

B2. Ask non-POEMB principal investigators whether they are ***aware*** of any important accomplishments of the programs that received the award and whether they or any of their colleagues ***maintained an interest*** in the application of the techniques of molecular biology in the areas of cardiovascular and/or pulmonary diseases.

Primary Codes	Secondary Codes	Tertiary Codes
AWARE Discussion of whether or not interviewee is **aware** of any important accomplishments on the part of the Programs that received the award. Only for non-POEMB principal investigators. [Sect. II, #'s 1-3]	**FIELD:** discussion of the important accomplishments in the application of molecular biology in the areas of heart and lung diseases. This refers to the field as a whole, and not just to accomplishments on the part of the Programs. [Sect. II, #1]	
	PROGRAMS: discussion of the important accomplishments on the part of the three Programs. [Sect. II, #3]	
	APART: when a disparity exists between the important accomplishments in the field and the accomplishments of the three Programs.	
MAINTAIN Discussion of whether or not interviewee and/or interviewee's colleagues have **maintained an interest** in the application of molecular biology techniques in the areas of cardiovascular and/or pulmonary diseases. Only for non-POEMB principal investigators. [Sect. III, #'s 1-4]	**HOW:** indication of how interviewee and/or colleagues have maintained an interest, i.e., the kinds of research they have done since the application. [Sect. III, #1]	
	FUNDING: indication of types and sources of funding for continuing research in molecular biology and heart and lung. Also, comparison between these sources and POEMB. [Sect. III, #1a]	
	PROPOSED: indication of whether researcher pursued research originally proposed in the POEMB application. [Sect. III, #1b]	
	INFLUENCE: discussion of whether the research and accomplishments of the POEMBs has influenced the interviewee. [Sect. III, #2]	
	NETWORKUN: indication of whether the interviewee collaborates or otherwise interacts with Program researchers on a scientific or intellectual level regarding issues of molecular biology and heart and lung research. [Sect. III, #3]	
	IMPETUS: whether and how POEMB served as an impetus to pursue molecular biological research in cardiovascular and pulmonary fields, on both individual and institutional levels.	

B2. Ask non-POEMB principal investigators whether they are ***aware*** of any important accomplishments of the programs that received the award and whether they or any of their colleagues ***maintained an interest*** in the application of the techniques of molecular biology in the areas of cardiovascular and/or pulmonary diseases.

Primary Codes	Secondary Codes	Tertiary Codes
NETWORK Discussion of whether a formal or informal network was created in which researchers associated with the three Programs could interact, collaborate, or otherwise share information and resources with one another. [Sect. IV, 1a-1c]		
RECOMMEND Discussion of interviewee's recommendations to NHLBI for the future of POEMB and the pursuit of the primary objectives. [Sect. V, #1]		

Battelle

. . . Putting Technology To Work

Centers for Public Health Research and Evaluation
2101 Wilson Boulevard
Arlington, Virginia 22201