Relationship Between Time of First Publication and Subsequent Publication Success Among Non-PhD Physician-Scientists

Kevin R. Riggs, MD, MPH
Zachary J. Reitman
Thelma J. Mielenz, PhD
Philip C. Goodman, MD

Abstract

Background Studies have shown that publication of work during medical school and residency is associated with higher numbers of later publications and citations of published research. However, it is unknown whether this association exists for non-PhD physician-scientists and whether the association persists later into their careers.

Methods We extracted publication records from the curricula vitae (CVs) of 102 corresponding authors of articles published in 2008 in the *New England Journal of Medicine* and *JAMA*, and obtained those authors' citation records from Web of Science. We used regression models to examine the association between time of first publication and later publication and citation rates for the entire postgraduate career and a recent 2-year period.

Results After adjusting for time since medical school graduation, sex, location of medical school (United States or not United States), and additional non-PhD degrees, we found that authors who first published before graduating from medical school had a greater mean number of publications after medical school and during the period from 2006 to 2007 (164 and 28, respectively) than those who first published during the 5 years afterward (111 and 19, respectively) and those who first published more than 5 years after graduation (59 and 13, respectively). Similarly, authors who first published before graduating from medical school had a greater mean number of citations of their published work since graduation and of publications from 2006 to 2007 (4634 and 333, respectively) than those who first published during the 5 years afterward (2936 and 183, respectively) and those who first published more than 5 years after graduation (1512 and 143, respectively).

Conclusions Early publication is associated with higher numbers of publications and more citations of published research among non-PhD physician-scientists. This association persists well into a researcher's career.

Background

Physician-scientists are defined as "individuals with an MD degree who perform medical research as their primary activity."[1] Experts have expressed concern over the past several decades about the decreasing supply of physician-scientists.[2,3] The response has been numerous initiatives, including new funding opportunities for research during medical school[4] and loan repayment options for graduates who enter research careers.[5] Targeting the pool of traditionally trained physicians is critical because, despite the combined MD-PhD track, designed to train physician-scientists, only a minority of physician-scientists have MD-PhD degrees.[1] Previous research has shown that participation in research during medical school or residency is associated with an increased likelihood of choosing an academic career.[6] Spending time on research during this critical time in medical training has drawbacks, the most obvious being less time to focus on clinical training.

Given the potential drawbacks of spending time on research projects during the formative clinical years, it is important to identify the benefits of such endeavors. Several previous studies that retrospectively examined cohorts from a single medical school or residency and compared the research output of those who did research or published during their training to those who did not

Kevin R. Riggs, MD, MPH, is a resident, Department of Medicine, Duke University Medical Center; **Zachary J. Reitman,** is an MD-PhD candidate, Duke University School of Medicine; **Thelma J. Mielenz, PhD,** is Assistant Professor, Department of Epidemiology, Mailman School of Public Health, Columbia University, and Adjunct Assistant Professor, Division of Physical Therapy, Department of Allied Health Sciences, University of North Carolina; **Philip C. Goodman, MD,** is Professor, Department of Radiology, Duke University Medical Center

Funding: The authors report no external funding source for this study.

Corresponding author: Kevin R. Riggs, Duke University Medical Center, Room 8254DN, 2301 Erwin Road, Durham, North Carolina 27710, kevin.riggs@duke.edu

Received March 23, 2011; revisions received: September 4, 2011, and December 1, 2011; accepted: December 7, 2011.

DOI: http://dx.doi.org/10.4300/JGME-D-11-00068.1

identified an association between early research and later success.[7,8] Because subjects in those studies self-selected whether to participate in early research, personal characteristics, such as innate interest in research, were likely to differ widely between those groups. To address this issue, we retrospectively sampled a group of active and successful researchers. Through this approach, we sought to remove at least some of the bias associated with prospective studies because, over time, personal interest in research has likely become comparable among the groups. We tested the hypothesis that early research is associated with future academic success among those students who go on to pursue careers as non-PhD physician-scientists, both in terms of quantity of research (number of publications) and quality of research (citations of published research). Additionally, by examining a recent 2-year period, we determined whether an association between early research and later success persists late into a researcher's career or attenuates over time.

What was known

Publication during medical school and residency is known to be associated with higher numbers of later scientific publications.

What is new

A study of publication records showed authors who first published before medical school graduation had higher average number of publications than authors who first published during the 5 years after graduation, and those who first published more than 5 years after graduation.

Limitations

Small sample. "Publication success" measured by publication in two journals, and journal choice may have introduced bias in favor of clinical research. The analysis was not able to account for other important contextual factors such as research mentoring.

Bottom line

Early publication is associated with higher publication success for non-PhD physician-scientists and that association persists later into these researchers' career.

Methods

Subjects

To identify non-PhD physician-scientists, we examined each issue of the *New England Journal of Medicine* (*NEJM*) and *Journal of the American Medical Association* (*JAMA*) for the year 2008. We e-mailed corresponding authors of Original Articles in *NEJM* and Original Contributions in *JAMA* who had an MD, MD equivalent (DO, MBBS, MB ChB, and other variations), or potential MD equivalent (FRACP, FRCP, and other) degrees, and no other doctoral level research degree (PhD, DPH, or other), and we asked them to submit a current curriculum vitae (CV). This study was determined to be exempt research by the Duke University Institutional Review Board, where the data were collected. The University of North Carolina institutional review board approved this study as existing or nonresearch data.

Extraction of Data from CVs

Two of the authors (KRR and ZJR) independently extracted data from the CVs. Data extracted included sex, year of graduation from medical school, whether medical school was in the United States, residency type, other postbaccalaureate-level degrees, number of publications published before or during the year of medical school graduation, number of publications published up to 5 years after the year of medical school graduation, number of publications published after the year of medical school graduation through the end of 2007, and number of publications published in 2006–2007.

We included letters, proceedings of meetings, and online publications for which the author was cited by name. We did not include entries that referred to nonwritten media (DVDs, CDs, and other forms), works that were only edited, or educational materials for which the author was not cited by name. We excluded abstracts; we did not include entries under "abstracts" or any header that included abstracts (ie, "abstracts/letters"). We also excluded entries from "miscellaneous" or "other," unless their contents were further described.

Extraction of Data from Web of Science

We searched Web of Science (WoS; Thomas Reuters, New York, NY) both for publications and total citations for the time periods of medical school graduation through 2007 and 2006 to 2007. Author attribution using electronic databases represents a significant gap in the field of information sciences, and much research is currently being conducted in the field to address it.[9] Because of this, we were unable to find a standard method for conducting a search using WoS. We chose to use a relatively simple method that maximized specificity (not attributing papers to the author of interest) rather than sensitivity (not missing papers that should have been attributed to the author of interest). We conducted each search of Science Citation Index Expanded (SCI-EXPANDED), using the researcher's name as it was written in the index article (last name, first initial, and middle initial if included). For women researchers, we checked the first publication on the CV to determine if there had been a name change. In cases where there was a name change, we searched both names. We then refined the search results by institution according to the institutions listed on the researcher's CV. These limits were intended to disambiguate authors who had common names. Finally, we bounded each search by the years of interest, starting either with the year of medical school graduation, for the researchers who had published before

medical school graduation, or the year of first publication according to the CV for those who first published after graduation.

After conducting the search on WoS, we compared the number of publications found to the number found on the CV. If the number on WoS was less than 50% of the number on the CV, we went back to the CV to look for any alternative name spellings that could explain the low yield. If the number on WoS was greater than 100% of the number on the CV, we assumed that there may be more than one author by the name searched at one of the institutions searched, and we attempted to exclude any papers that did not belong to the researcher of interest.

Data Analysis

The 4 outcome variables were (1) number of publications from medical school graduation through 2007 as determined from the CV; (2) number of citations of the publications from graduation through 2007 as determined by WoS; (3) number of publications during 2006 to 2007 as determined from the CV; and (4) number of citations to the 2006 to 2007 publications as determined by WoS. The main exposure variable was time of first publication, as defined in 3 categories: "early publishers" (first publication on the CV was during or before the year of medical school graduation), "middle publishers" (first publication on the CV was between 1 and 5 years after the year of medical school graduation), and "late publishers" (first publication on the CV was more than 5 years after the year of medical school graduation).

To help identify potential confounders prior to modeling, we used bivariate analyses to examine whether any of the independent variables were unequally distributed across time to publication and to examine the relationship between the 4 outcomes and each of the independent variables. We also evaluated linearity between continuous variables and outcome measures as well as collinearity between each of the categorical variables prior to modeling.

For each of the 4 main outcome measures, we used a multiple linear regression analysis of covariance (ANCOVA) model to estimate the mean number of publications or citations, adjusted for covariates. We treated the 3-category variable for time of first publication as nominal and created indicator variables, with early publishers as the reference category. For the publications from medical school graduation through 2007 and citations to those publications, we adjusted models for years since medical school graduation (as of the end of 2007), sex, whether the researcher went to medical school in the United States or not, and presence of an additional graduate degree (none, MPH, or other masters degree). For publications during 2006 to 2007 and citations to those publications, we adjusted the models for the same variables, except that we used stage of career, a 4-category nominal variable (0–8, 9–16, 16–24, or more than 24 years since graduation), rather than years since medical school as a continuous variable because of the linearity assumption. We did not include residency type in any of our analyses because the large percentage of participants who specialized in internal medicine (including subspecialties) meant that the remaining residency types had too few entries for a meaningful regression model.

We initially included all potential covariates (except for residency type) in the model. The model was then reduced by sequentially removing variables that did not change the number of publications or citation estimates for each category of time of first publication a meaningful amount. It appeared that for most outcome models, time since medical school graduation (career stage in 2 of the models) was the most important confounder and that sex and the location of medical school were also weak confounders. Whether the researchers had an additional masters degree was not a confounder in any of our models. In the end, we decided to leave all of the variables in the model as we felt them to be potentially important confounders.

Results

Of the 199 researchers we e-mailed, 102 (51%) sent a CV. Characteristics of responders and nonresponders are shown in TABLE 1. Those who responded were more likely to have a correspondence address in the United States ($P = .04$). No other differences were statistically significant.

TABLE 1 **Characteristics of Responders and Non-Responders to CV Request**[a]

Characteristic	% of Responders (n = 102)	% of Non-Responders (n = 97)
Address in United States[b]	74	60
RTC	44	42
First Author	72	71
Last Author	23	24
Non-PhD Postgraduate Degree		
MPH	12	11
Other	14	10

Abbreviations: RTC, randomized controlled trial.

[a] Significance tests for comparisons between groups was based on Pearson's χ^2 test.

[b] $P < .04$.

TABLE 2 **Baseline Characteristics of Physician-Scientists by the Time of First Publication**

Characteristic	Overall Mean % (±SD) (n = 83)[a]	Mean % of Early Publishers (±SD) (n = 24)	Mean % of Middle Publishers (±SD) (n = 42)	Mean % of Late Publishers (±SD) (n = 17)
Male[b]	76	92	74	59
Medical School in US	69	83	64	59
Years Since Medical School Graduation	20.7 (9.7)	17.8 (10.6)	21.6 (9.4)	22.8 (8.9)
Non-PhD Postgraduate Degrees				
MPH Degree	16	13	14	24
Other	29	25	29	35
Specialty				
Internal Medicine	57	63	62	35
Surgery	6	4	10	0
Pediatrics	12	8	10	24
Obstetrics and Gynecology	6	4	5	12
Other	19	21	14	29

[a] Overall significance tests based on one-way ANOVA for continuous characteristics and Pearson's χ^2 test for categorical characteristics.
[b] $P < .048$.

Of the 102 CVs we received, 83 were included in the final analysis. Reasons for exclusion (and number of CVs excluded) were the researcher had a PhD (6); the CV contained no publication list (4) or the list was marked as "selected publications" (3); there was insufficient information on the CV (3); the CV was not updated through the end of 2007 (2); and the CV was sent as an unopenable file (1). In cases where publication lists were absent or incomplete, we would have been unable to accurately categorize researchers as early, middle, or late publishers, which was a critical first step.

Characteristics of early, middle, and late publishers who were included in the final analysis are shown in TABLE 2. Notably, early publishers were more likely to be men and had on average postmedical school careers approximately 4 years shorter than the middle or late publishers.

Early publishers had a mean number of publications before medical school graduation of 3.2 (range, 1–25) and 5.3 during the first 5 years after graduation (range, 0–33). Middle publishers had a mean number of publications during the first 5 years after graduation of 3.2 (range, 1–12). By definition, middle publishers had no publications before medical school graduation, and late publishers had no publications before medical school graduation or during the first 5 years after graduation.

Results of unadjusted and adjusted analyses are shown in TABLE 3. In all 4 outcomes, after adjusting for covariates, we found early publishers had 47% to 82% more publications/citations than middle publishers and 116% to 206% more publications/citations than late publishers. For the publications after medical school graduation through 2007, the citations of these publications, and publications in 2006–2007, the magnitude of each of these mean differences was also statistically significant. For the citations of publications in 2006 to 2007, the same trend for the early publishers continued but these differences were not significant.

Discussion

We show here a positive relationship between early publishing and subsequent publication success among a select group of non-PhD physician-scientists. Those who published before graduation from medical school published the most and had the most citations of their work in their postmedical school careers. They were followed by those who first published within 5 years of medical school graduation and finally by those who first published even later.

The relationship between early publication and later publication success likely does not represent bias introduced by the study design. The late publishing group, by

TABLE 3 **Unadjusted and Adjusted Mean Number of Publications/Citations by the Time of First Publication**

Time to Publication	No. of Publications After Medical School[a]	No. of Citations to Publications After Medical School[a]	No. of Publications in 2006–2007[b]	No. of Citations to Publications in 2006–2007[b]
Unadjusted Mean No.				
Early Publishers	137	3925	25	299
Middle Publishers	119	3177	20	195
Late Publishers	77	1919	15	161
Adjusted Mean No.[c]				
Early Publishers	164	4634	28	333
Middle Publishers	111	2936	19	183
Late Publishers	59	1512	13	143

[a] Based on the beta estimates from a multiple linear regression model, adjusted for years since medical school graduation, sex, location of medical school (United States or not United Stages), and additional degrees (none, MPH, or other masters).
[b] Based on the beta estimates from a multiple linear regression model, adjusted for stage of career, gender, location of medical school (United States or not United States), and additional degrees (none, MPH, or other masters).
[c] Overall *P* value (2 *df*) of <.05 for each adjusted outcome, except for citations of publications between 2006 and 2007 ($P = .145$).

definition, had no publications during the first 5 years after medical school graduation, while early and middle publishers had a "head start" of a mean of 5.3 and 3.2 publications, respectively, during that time. However, these differences are small compared to the differences in career publications. Additionally, we designed the study to address this potential bias by also analyzing total publications and citations during the 2006 to 2007 time period. We found similar trends between groups for publications in 2006 to 2007 and total publications (TABLE 3). This similar trend in both time periods added assurance that a real difference exists between the groups.

Our study has several limitations. First, the small sample size, stringent inclusion criteria, and moderate participation rate make it difficult to generalize the results to all non-PhD physician-scientists. Second, we did not attempt to differentiate between clinical and basic science researchers, and the value of early research may be different for those who choose these very different careers. Third, our approach of extracting data from CVs and from WoS introduced some potential for error. However, the alternative of asking researchers to complete detailed surveys regarding all their publications would likely have led to lower participation than simply asking them to send their CVs. Fourth, we were unable to include residency type in our regression models due to the small number of participants from several of the major specialties. The fairly even distribution of residency types among the 3 groups makes it unlikely that it would have been a major confounder, but it may be more difficult to generalize our findings to non-internal medicine specialties. Fifth, there may be important predictors to research success, such as quality of early mentorship, which are difficult to capture in a study such as this and are thus difficult to control for. Sixth, we chose publication and citation counts as surrogates for research career success, which may not be the most important factors for success. Other outcomes, such as achieving tenure or securing grants, could also be measured. However, information for these outcomes is more difficult to obtain, and the "value" of those outcomes is likely to vary among institutions. Finally, our study design does not allow reporting a causal relationship, which would likely require a difficult-to-administer randomized trial.

While we have demonstrated a relationship between early publication and later publication and citation rate, more work is warranted to test for a causal effect between early research and future research success. This relationship could reflect an association of certain individual characteristics with both early publishing and later high publication rates. However, this relationship could also be explained by a causal relationship between early research and later publication success. These possibilities are not exclusive.

Conclusions

Among a highly selected group of non-PhD physician-scientists, early publication was associated with more publications and more citations of published research, and this association persisted well into a researcher's career. While this research does not prove causal effect, the association alone may be enough to influence trainees considering undertaking research projects, as well as medical

schools and postgraduate training programs considering offering research opportunities. This study adds weight to the argument that early research is an important component of the education of future researchers.

References

1 Ley TJ, Rosenberg LE. The physician-scientist career pipeline in 2005: build it, and they will come. *JAMA*. 2005;294(11):1343–1351.

2 Wyngaarden JB. The clinical investigator as an endangered species. *N Engl J Med*. 1979;301(23):1254–1259.

3 Rosenberg L. Physician-scientists—endangered and essential. *Science*. 1999;283(5400):331–332.

4 Gallin EK, Le Blancq SM. Launching a new fellowship for medical students: the first years of the Doris Duke Clinical Research Fellowship Program. *J Investig Med*. 2005;53(2):73–81.

5 Ley TJ, Rosenberg LE. Removing career obstacles for young physician-scientists—loan-repayment programs. *N Engl J Med*. 2002;346(5):368–372.

6 Straus SE, Straus C, Tzanetos K. Career choice in academic medicine: systematic review. *J Gen Intern Med*. 2006;21(12):1222–1229.

7 Brancati FL, Mead LA, Levine DM, et al. Early predictors of career achievement in academic medicine. *JAMA*. 1992;267(10):1372–1376.

8 Dorsey ER, Raphael BA, Balcer LJ, Galetta SL. Predictors of future publication record and academic rank in a cohort of neurology residents. *Neurology*. 2006;67(8):1335–1337.

9 Smalheiser NR, Torvik VI. Author name disambiguation. In: Cronin B, ed. *Annual Review of Information Science and Technology*. Vol 43. Medford, NJ: Information Today; 2009:287–313.