

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

**The GRE Over the Entire Range of Scores Lacks Predictive Ability for PhD Outcomes  
In the Biomedical Sciences**

**Linda Sealy<sup>1,2\*</sup>, Christina Saunders<sup>3</sup>, Jeffrey Blume<sup>3</sup> and Roger Chalkley<sup>1,2</sup>**

<sup>1</sup>Office of the Dean of Basic Sciences, School of Medicine, Vanderbilt University,  
Nashville, Tennessee, United States of America

<sup>2</sup>Department of Molecular Physiology and Biophysics, School of Medicine, Vanderbilt  
University, Nashville, Tennessee, United States of America

<sup>3</sup>Department of Biostatistics, School of Medicine, Vanderbilt University,  
Nashville, Tennessee, United States of America

\*corresponding author

E-mail: [linda.sealy@vanderbilt.edu](mailto:linda.sealy@vanderbilt.edu) (LS)

## 44 **Abstract**

45 The association between GRE scores and academic success in graduate programs is currently of  
46 national interest. GRE scores are often assumed to be predictive of student success in graduate  
47 school. However, we found no such association in admission data from Vanderbilt's Initiative  
48 for Maximizing Student Diversity (IMSD), which recruited historically underrepresented  
49 students for graduate study in the biomedical sciences at Vanderbilt University spanning a wide  
50 range of GRE scores. This study avoids the typical biases of most GRE investigations of  
51 performance where only high-achievers on the GRE were admitted. GRE scores, while collected  
52 at admission, were not used or consulted for admission decisions and comprise the full range of  
53 percentiles, from 1% to 91%. We report on the 29 students recruited to the Vanderbilt IMSD  
54 from 2007-2011 who have completed the program at this date. While the data set is not large, the  
55 predictive trends between GRE and long-term graduate outcomes (publications, first author  
56 publications, time to degree, predoctoral fellowship awards, and faculty evaluations) are  
57 remarkably null and there is sufficient precision to rule out even mild relationships between GRE  
58 and these outcomes. Career outcomes are encouraging; many students are in postdocs, and the  
59 rest are in stage-appropriate career environments for such a cohort, including tenure track  
60 faculty, biotech and entrepreneurship careers.

61

62

63

64

65

## 66 **Introduction**

67           Recently Moneta-Kohler et al. [1] published a detailed statistical analysis of the lack of  
68 ability of the GRE to predict performance in graduate school in the biomedical research arena at  
69 Vanderbilt. A similar study was published by Hall et al. [2] from the University of North  
70 Carolina Chapel Hill. However, there was a limitation to the overall conclusions in that the  
71 range of GRE scores did not cover scores lower than approximately 50%. In order to test if such  
72 a limitation impacted the predictive ability of the GRE, we would need to admit students for  
73 whom we had GRE information, but where the admitted students covered the entire range of  
74 scores with no bias or cut-off (deliberate or otherwise) in the level of the score. This is a difficult  
75 requirement, as admissions committees normally do not pursue applicants with very low GRE  
76 scores, even if other aspects of the application might appear to be competitive. We are aware that  
77 a fairly significant number of schools are electing to not use GRE scores at all in making  
78 admissions decisions [3]. Other schools may be considering whether or not to require GRE  
79 scores, but have not yet taken action. All of these schools surely might benefit if there were to  
80 be an experiment in which we assayed the predictive ability of the GRE scores over the entire  
81 range of scores.

82

83           We report that we have performed this natural experiment with GRE scores covering the  
84 range from 1st to 91st percentile, in an approach where the scores, although submitted as part of  
85 the application, were not considered in the selection of incoming graduate students. This came  
86 about in the following way. In 2007 Vanderbilt was awarded an NIGMS-funded IMSD program  
87 with the goal of increasing the number of students from underrepresented groups completing  
88 PhDs in the biomedical sciences. This program was a redesign of our previous IMSD post

89    baccalaureate program in response to the NIH stipulation in 2006 that students in the program  
90    had to be matriculated as graduate students, not post baccalaureates. We were aware that  
91    increasing the number of historically underrepresented (UR) students in our PhD programs might  
92    result in another school(s) not enrolling these students, and the overall pool of UR PhD trainees  
93    would remain static. This was because at that time the pool of high qualified UR students, when  
94    quantitative metrics (GRE, GPA) were a key driver of the assessment, was in insufficient supply.  
95    The authors had already collected data over a ten-year period indicating that for  
96    underrepresented students at least the GRE at the levels usually expected for admission offered  
97    no guidance in terms of achievements of long term PhD training goals. Consequently, we  
98    decided that removing the barrier of GRE scores to admission would actually lead to an overall  
99    increase in historically underrepresented PhD trainees.

100

101        Therefore, in 2007 the Vanderbilt IMSD program adopted a fully holistic approach to  
102    admissions. The GRE scores were recorded as a required part of the application process, but  
103    they were essentially ignored by the IMSD admissions committee, which operated in a separate  
104    fashion from our regular interdisciplinary graduate program (IGP) admission committee. This  
105    resulted in a group of students who were eligible for IMSD support (as defined by NIGMS)  
106    admitted with GRE scores over the full range (1-90% GRE-V and 11-91% GRE-Q). Details of  
107    the fully holistic approach are presented below, but relied heavily on letters of recommendation,  
108    personal statements and interviews. If these factors were strong, no GRE score was too low to be  
109    admitted.

110

111 Over the next four years, 29 students, were admitted in this fashion, all of whom have  
112 now completed the PhD graduation cycle, so we are able to evaluate the outcomes of admissions  
113 strategies which cover the entire GRE range (including both very high and low scores), under  
114 conditions in which the admissions process operated obliviously to the scores themselves. The  
115 measures we have used to evaluate outcomes performance in biomedical research were also used  
116 in our previous report [1] on the lack of predictive value of the GRE. These include: number of  
117 first or other order author papers, receiving competitive fellowship awards, time to degree, a  
118 detailed faculty evaluation at the time of graduation, and an initial review of career development  
119 in scientific areas. Of the 29 URM students who participated in the IMSD program over this  
120 time period, 27 have now graduated (25 PhD, 2 MS) with two students dropping out early as a  
121 consequence of health problems. We present here the outcomes of the 25 PhD graduates and the  
122 relationship of these outcomes to their GRE scores.

123

## 124 **Materials and Methods**

125 GRE (Quantitative and Verbal) scores and academic performance data from 25 IMSD students  
126 who matriculated from 2007 to 2011 were collected and examined. Academic performance  
127 outcomes of interest were: time elapsed in program (i.e., months to degree), number of  
128 publications, number of first-author publications, fellowship status (any or F31), Vanderbilt  
129 faculty ranking (10 = best, 50 = worst). Table 2 provides univariate summaries (e.g., mean,  
130 median, standard deviation, inter-quartile range) of these variables. Figure 2 presents histograms  
131 of the continuous outcome variables. Regression modeling was used to assess the degree of  
132 association between GRE outcomes and academic outcomes. Specifically, Poisson regression  
133 was used to model publication counts (accounting for length of time in the program), months to

134 degree, and faculty ranking. Logistic regression was used to model receipt of fellowship. For all  
135 models, we report point estimates, model robust standard errors, and 95% confidence intervals  
136 (CIs). We plot each performance measure as a function of GRE scores and include the fitted  
137 regression line as well as a locally weighted scatterplot smoother (lowess) line to visually assess  
138 linearity assumptions and model fit. Confidence intervals were plotted to demonstrate the degree  
139 of precision afforded by the data at the 95% level. Any relationship between GRE scores and  
140 outcomes would be captured in the slope of these regression lines. While it is not possible to  
141 prove the null hypothesis that GRE scores and outcomes are not related, it is possible to provide  
142 an upper bound on the largest potential association. The 95% CI provide this bound and  
143 comprise the set of associations supported by the data. As we will see from the data, despite the  
144 small sample size, these CIs do not support mild or strong associations between GRE scores and  
145 outcomes. For a sensitivity analysis, we compared academic outcomes between the first quartile  
146 and the fourth quartile of GRE scores. If any association were present, such an analysis should at  
147 least yield exaggerated point estimates of the association effect. The research was approved by  
148 Vanderbilt University IRB (151678). Consent was not given as data were analyzed  
149 anonymously.

150

## 151 **Results**

152 In Fig 1 we report the range of GRE scores among the 27 URM students admitted into  
153 the graduate program in the biomedical sciences at Vanderbilt from 2007 through 2011 and who  
154 completed a PhD or MS degree. The admission decisions for these students during this time  
155 period was determined by the IMSD admissions/advisory committee, and although the student's  
156 GRE score was recorded in our databases, it has only been used for outcomes studies long after

157 the admissions event. The range of GRE scores among this group of students varies across the  
158 spectrum for students who were admitted in response to a detailed analysis of the committee's  
159 assessment of the likelihood of the student's success in research. The committee's assessment  
160 was based primarily on the non-quantitative components of the application, including a close  
161 reading of the letters of recommendation and the student's personal statement. The student's  
162 transcript was evaluated, primarily to assess adequate coursework preparation for biomedical  
163 PhD coursework. A wide range of GPAs were accepted. We sought to place the overall and  
164 science GPAs in the context of the college or university and the life events of the applicant. For  
165 example, students with extensive work and/or family responsibilities might reasonably be  
166 expected to end up with lower GPAs due to time demands. The lowest GPA accepted among  
167 this group of students was 1.8. Finally, all students were invited to campus for an interview visit  
168 that was also given significant consideration.

169

170 **Fig 1. GRE Quantitative and Verbal Scores of IMSD Students Matriculating from 2007-11**  
171 **who completed a PhD or Master's degree.** Top panel depicts GRE-Q% and lower panel depicts GRE-  
172 V% for 27 students who completed either a PhD (blue symbols) or Master's degree (red symbols).

173

174 From the 27 students with GRE scores shown in Fig 1, 93% have graduated with the PhD  
175 and 7% left with an MS degree. Of the 25 students who completed the PhD, 84% continued to  
176 postdoctoral positions. Four students did not continue on to postdocs, choosing instead to move  
177 to industry, consulting, medical school, or an academic faculty position. Overall, the outcomes of  
178 this cadre of GRE-blind admitted students are strikingly parallel to those of students admitted  
179 through the traditional route (using much higher GRE scores) over the same time period [4]. As  
180 indicated in Fig 1, we have a wide range of GRE scores among this group. This unusual group

181 provided us with a means to test the predictive value of GRE scores over a much wider range  
182 than most admissions committees will typically tolerate.

183

184 Fig 2 shows histograms of the data analyzed in this study: range and frequencies of GRE  
185 scores, number of publications, number of first author publications, months to degree, and  
186 faculty ranking. The faculty ranking is obtained upon the student's completion of their Ph.D. The  
187 ranking is comprised of the sum of scores for each of ten questions, listed in Table 1. The  
188 questions cover a range of areas that are often informally assessed as measures of developing  
189 into a successful, independent scientist; many would fall into the area of the social/emotional  
190 learning skillset. We ask the PhD faculty mentor to score their newly-minted PhD student from  
191 one to five, with one being best. Thus, the top ranking possible is a 10, if the student received a  
192 score of one for each of the ten questions. Student rankings ranged from 12 to 39 with a median  
193 of 21.5. The other metrics are self-explanatory, with number of publications ranging from one to  
194 thirteen (median= 4) and first author publications from one to six (median=2). Note that students  
195 are expected to publish at least one first author paper as a requirement for the PhD in most of our  
196 PhD granting programs. The time to degree for these students ranged from slightly more than 4  
197 years, to just over 7 years (median = 5.66 years). In addition to the data shown in Fig 2, we also  
198 included whether or not the student obtained an individual fellowship in a national competition  
199 (F31, AHA, DOD, etc) as an additional metric (see Figs 6 and 7). Summary statistics of the data  
200 for this study are presented in Table 2. The hypothesis we test is that GRE scores are associated  
201 with future performance in a biomedical graduate program. This association will be measured by  
202 the slope in a regression model, to be explained shortly.

203



204 **Fig 2. Histograms of outcomes data.** Frequencies of GRE scores (Q% and V%), number of publications,  
205 months to degree, and faculty ranking are shown as indicated.

206

207 **Table 1. Faculty rating of student at exit.**

1. Ability to handle classwork needed for your PhD program
2. Drive and determination
3. Creativity and imagination in terms of experimental design and interpretation
4. Technical ability
5. Keeping up with the literature
6. Output – translating observations into a presentable paper
7. Ability to write creatively
8. Leadership in the lab and department
9. Trajectory
10. Overall assessment as a productive scientist

208 Faculty mentors were asked to score student upon PhD completion using a scale of 1-5 as follows:  
209 1-outstanding; 2-excellent; 3-good; 4-fair; 5-poor. The faculty rating is the sum of the scores for 10  
210 questions.

211

212

213

214

215

216

217

218 **Table 2. Summary statistics of GRE data**

219

220

	N				
GRE-Quantitative	25	28	44	59	(45 ±21)
GRE-Verbal	25	22	43	60	(42 ±25)
No. of publications	25	3.0	4.0	7.0	(4.9 ±2.5)
No. of first author pubs	25	1.0	2.0	2.0	(2.0 ±1.1)
Any fellowship	25				
0					52% (13)
1					48% (12)
F31 fellowship	25				
0					68% (17)
1					32% ( 8)
Other fellowship	25				
0					84% (21)
1					16% ( 4)
Faculty ranking	20	17.8	21.5	30.5	(23.8 ± 8.1)
Months to degree	25	63.6	68.0	75.3	(68.8 ± 9.6)

228

229

230

231

232

*a b c* represent the lower quartile *a*, the median *b*, and the upper quartile *c* for continuous variables.  $\bar{x} \pm s$  represents  $\bar{X} \pm 1$  standard deviation. *N* is the number of non-missing values. Numbers after percents are frequencies.

233

234 **Lack of association between GRE scores and publications.**

235

236

237

238

239

240

241

242

We modeled the relationship between total number of publications and GRE scores using Poisson regression in Fig 3 for GRE-Q (left panel) and GRE-V (right panel). Solid curves show the fitted values from the regression models (dashed lines are 95% confidence intervals) and the grey curves show lowess smoothers (locally weighted scatterplot smoother). Increasing a student's GRE-Q score by 20 points increases their expected publication rate by just 3% (rate ratio = 1.03 with 95% CI 0.85 to 1.26). For instance, students with GRE-Q scores of 40 and 60 are expected to have 4.79 and 4.94 publications, respectively. Interesting, increasing a student's GRE-V score by 20 points decreases their expected publication rate by 4% (0.77, 1.19). For

243 instance, students with GRE-V scores of 40 and 60 are expected to have 4.83 and 4.62  
244 publications, respectively.

245

246 **Fig 3. Associations between quantitative and verbal GRE scores and total number of**  
247 **publications.** Solid curves show the fitted values from the Poisson regression models (dashed lines are 95%  
248 confidence intervals) and the grey curves show lowess smoothers (locally weighted scatterplot smoother).

249

250 We do not judge these minor differences to be significant, although the same negative  
251 correlation with GRE-V scores was also observed when the total number of first author  
252 publications and GRE scores was modeled using Poisson regression in Fig 4. Increasing a  
253 student's GRE-V score by 20 points decreases their expected publication rate by 13% (rate ratio  
254 =0.87 with 95% CI 0.69 to 1.11). For instance, students with GRE-V scores of 40 and 60 are  
255 expected to have 1.97 and 1.72 first author publications, respectively. Increasing a student's GRE-  
256 Q score by 20 points increases their expected first author publication rate by 8% (rate ratio =  
257 1.077 with 95% CI 0.88 to 1.32). For instance, students with GRE-Q scores of 40 and 60 are  
258 expected to have 1.94 and 2.09 first author publications, respectively, which is essentially no  
259 difference. We conclude that even when GRE scores below 20 percentile are in the mix,  
260 productivity as measured by the key currency of the scientific enterprise, namely publications -  
261 exhibits very little dependence, if any, on GRE scores.

262

263 **Fig 4. Associations between quantitative and verbal GRE scores and total number of first**  
264 **author publications.** Solid curves show the fitted values from the Poisson regression models (dashed lines are  
265 95% confidence intervals) and the grey curves show lowess smoothers (locally weighted scatterplot smoother).

266

## 267 **Lack of association between GRE scores and time to degree**

268           In Fig 5, months to degree is plotted vs either GRE-Q (left panel) or GRE-V (right  
269 panel). Again, the solid curve shows the fitted values from the Poisson regression model  
270 (dashed lines are 95% confidence intervals) and the grey curve shows a lowess smoother (locally  
271 weighted scatterplot smoother). We observe only a very minor correlation between higher GRE  
272 scores and shorter time to degree. Increasing either the GRE-Q or GRE-V by 20 percentage  
273 points leads to a decrease in expected time to degree attainment of 3% (rate ratio = 0.97 with  
274 95% CI 0.92 to 1.01) or 2% (rate ratio = 0.98 with 95% CI 0.93 to 1.03), respectively. This  
275 means that students with GRE-Q scores of 40 and 60 are expected to take 69 months and 67  
276 months to complete their degree, respectively. Likewise, students with GRE-V scores of 40 and  
277 60 are expected to take 69 months and 68 months to complete their degree, respectively.

278

### 279 **Fig 5. Associations between quantitative and verbal GRE scores and months to degree.**

280 Solid curves show the fitted values from the Poisson regression models (dashed lines are 95% confidence intervals)  
281 and the grey curves show lowess smoothers (locally weighted scatterplot smoother).

282

## 283 **Lack of association between GRE scores and fellowships.**

284           We are well aware that counting papers, either first author or total, has limitations –  
285 especially since neither metric captures the quality and/or impact of the publications. Such  
286 parameters are difficult to uniformly measure because they are often very field-specific, and  
287 sometimes the impact of research is not fully appreciated for years to come. Therefore, we  
288 sought to include individual fellowships obtained as one metric of student quality. We included  
289 fellowships that are reviewed nationally by panels of experts, providing a comparison between

290 students in this cohort against students at similar stages of training from other institutions around  
291 the country. Predoctoral fellowships obtained by this cohort are included in Table 3.

292

293 **Table 3: Individual Fellowships awarded to IMSD students**  
294 **matriculating from 2007-2011**

295

<b>Fellowship type</b>	<b>number awarded</b>
F31 Ruth L. Kirschstein National Research Service Award (NRSA) Predoctoral Fellowship	8
American Heart Association Predoctoral Fellowship	1
National Science Foundation Graduate Research Fellowship	1
UNCF Merck Graduate Science Research Dissertation Fellowship	1
Department of Defense Prostate Cancer Research Program Predoctoral Fellowship	1

296

297 Boxplots of GRE scores stratified by whether or not students received a fellowship are  
298 shown in Fig 6. From bottom to top, the horizontal lines of a boxplot show the min, 25th  
299 percentile, median, 75th percentile, and max values in a given group. In Fig 7 the predicted  
300 probability of obtaining a fellowship as a function of GRE score is presented. Interestingly,  
301 increasing a student's GRE-Q score by 20 points decreases their odds of receiving a fellowship  
302 by 35% (odds ratio = 0.65 with 95% CI 0.27 to 1.58). For instance, the predicted probability of  
303 receiving a fellowship for students with GRE-Q scores of 40 and 60 are 50% (95% CI 30% to  
304 71%) and 40% (95% CI is 15% to 64%), respectively. Alternatively, increasing a student's GRE-  
305 V score by 20 points increases their odds of receiving a fellowship by just 4% (odds ratio = 1.04

306 with 95% CI 0.55 to 1.98). The predicted probability of receiving a fellowship for students with  
307 GRE-V scores of 40 vs. 60 are 48% (28%, 68%) and 49% (25%, 73%), respectively. We  
308 conclude for this data set, that GRE scores have little value in predicting who will receive a  
309 fellowship; in fact, for GRE-Q we observed a negative correlation.

310 **Fig 6. Boxplots of GRE scores stratified by whether or not the students received a**

311 **fellowship.** Data are for fellowships listed in Table 3. The raw data points are overlaid. From bottom to top, the  
312 horizontal lines of a boxplot show the minimum GRE score, 25<sup>th</sup> percentile, median, 75<sup>th</sup> percentile, and max values  
313 in a given group.

314 **Fig 7. Predicted probability of obtaining a fellowship as a function of GRE scores.** Data are

315 for fellowships listed in Table 3. The raw data points are overlaid (0 = No fellowship, 1= Fellowship).

316

317 **Lack of association between GRE scores and faculty evaluation.**

318 At the completion of their doctoral training, each faculty mentor is asked to evaluate their  
319 PhD student on each of ten questions provided in Table 1. The student is not aware that they are  
320 or have been evaluated, and the evaluation is never shared with the student nor used for any other  
321 purpose. It is important to note that a lower ranking indicates a better evaluation, with 10 being  
322 the highest score possible and 50 the lowest score. Fig 8 (left panel) shows the association  
323 between GRE-Q score and faculty ranking. As in the prior figures, solid curves show the fitted  
324 values from the Poisson regression models (dashed lines are 95% confidence intervals) and the  
325 grey lines show the lowess curves. Corresponding data for GRE-V score and faculty ranking are  
326 presented in Fig 8, right panel. In each case the associations were small and actually negative  
327 (that is, higher GRE scores were associated with lower faculty rankings). Increasing GRE-Q by

328 20 points increases (worsens) the expected ranking by 5% (rate ratio = 1.05 with 95% CI 0.92 to  
329 1.19). For instance, for students with GRE-Q scores of 40 and 60, the expected rankings are  
330 23.56 and 24.65. Likewise, increasing GRE-V by 20 points increases (worsens) the expected  
331 ranking by 10% (rate ratio = 1.10 with 95% CI 1.003 to 1.217). For students with GRE-V scores  
332 of 40 and 60, the expected rankings are 23.57 and 26.04.

333

334 **Fig 8. Associations between GRE scores and faculty ranking.** Solid curves show the fitted values  
335 from the Poisson regression models (dashed lines are 95% confidence intervals) and the grey curves shows lowess  
336 smoothers (locally weighted scatterplot smoother).

337

338 The data indicate that GRE scores across the entire range of values in this cohort are not  
339 predictive of the outcome measures we assessed. We took one final approach – testing for  
340 differences in performance measures between the lower and upper quartiles of the GRE scores.  
341 To be clear, we compared students with very low scores (<25% GRE-Q or V) to students with  
342 very high scores (>75% GRE-Q or V). Although this approach does not use all the data, it  
343 would be expected to yield an upwardly biased estimate of the GRE outcome association. The  
344 results of such an analysis are shown in Table 4 (for GRE-Q) and Table 5 (for GRE-V). For both  
345 tables the first two columns show the mean and standard deviation (SD) of performance  
346 measures (number of publications, number of first author publications, months to degree, and  
347 faculty ranking) among students in the lower 25th percentile and the upper 25th percentile of  
348 GRE score. The third and fourth columns show the difference in mean performance measures  
349 between the lower and upper quartiles and the 95% confidence intervals. We see that the point  
350 estimates are modest at best, and all confidence intervals include zero as expected. Therefore,  
351 even when comparing very low scores, (a range that many graduate schools rarely admit

352 students) to high scores, we do not find evidence that a relationship exists even between the two  
353 most likely classes of students.

354

355 **Table 4: Mean (SD) of variables in lower and upper quartiles of GRE-Q**  
356 **and 95% CIs for their difference**

	Lower Q	Upper Q	Difference	95% CI
No. of publications	4.6 (2.2)	4.3 (2.1)	0.24	(-2.39, 2.86)
No. of first author pubs	1.7 (0.8)	1.5 (0.8)	0.21	(-0.77, 1.2)
Months to degree	75 (7.6)	69 (7)	5.95	(-2.97, 14.88)
Faculty ranking	21.5 (9.3)	27 (7.7)	-5.5	(-18.15, 7.15)

357 The first two columns show the mean and standard deviation (SD) of performance measures among students in the  
358 lower 25<sup>th</sup> percentile and the upper 25<sup>th</sup> percentile. The third and fourth columns show the difference in mean  
359 performance measures between the lower and upper quartiles and the 95% confidence intervals.

360

361 **Table 5: Mean (SD) of variables in lower and upper quartiles of GRE-V**  
362 **and 95% CIs for their difference**

	Lower Q	Upper Q	Difference	95% CI
No. of publications	5.3 (3.5)	4.7 (2.3)	0.62	(-2.96, 4.2)
No. of first author pubs	2.7 (1.7)	1.7 (0.8)	1.05	(-0.6, 2.69)
Months to degree	72.2 (11.1)	67.4 (9.6)	4.86	(-7.8, 17.52)
Faculty ranking	20.8 (5.2)	27.2 (9.2)	-6.37	(-17.67, 4.93)

363 The first two columns show the mean and standard deviation (SD) of performance measures among students in the  
364 lower 25<sup>th</sup> percentile and the upper 25<sup>th</sup> percentile. The third and fourth columns show the difference in mean  
365 performance measures between the lower and upper quartiles and the 95% confidence intervals.

366

## 367 **Outcomes of the cohort to date**

368 For the 25 students in the cohort analyzed here, the final question we can ask is where are  
369 they now? As mentioned earlier, most of the cohort moved on to a postdoctoral position upon  
370 PhD completion at a range of research intensive institutions listed in Table 6. The students in



371 this cohort completed their PhDs between spring 2012 and summer 2017, so some have had time  
372 to move to a position beyond the first postdoc. So far after completing their first postdoctoral  
373 position, two individuals have moved on to Biopharma, one who is developing a start-up  
374 company, one moved to an administrative position at NIH, and one is now a tenure-track  
375 assistant professor. At of the time of this writing (June 2018), none of this cohort of 25 students  
376 have left science.

377

378 **Table 6: Postdoctoral institutions for IMSD students upon PhD completion**

Harvard University
Mt Sinai Icahn School of Medicine
University of Texas Health Science Center
Northwestern University
Yale University
University of Florida
Vanderbilt University Medical Center
John Hopkins University
National Institutes of Health
Baylor College of Medicine
University of Colorado
Michigan State University
St Jude Children's Research Hospital
Case Western University
University of Washington

Charles R. Drew University of Medicine and Science
Vanderbilt University

379 Institutions where IMSD students who matriculated from 2007-2011  
380 completed first postdocs  
381

382

## 383 **Discussion**

384 As a result of the admissions process adopted by the Vanderbilt IMSD program over a  
385 decade ago, we now have a cohort of graduate students whose GRE scores spanned the entire  
386 range from 1-91 percentile who have completed the PhD. This analysis includes 25 IMSD  
387 students who matriculated into our biomedical research programs from 2007-2011 and  
388 completed PhDs beginning in 2012 to summer 2017. We consistently observed only associations  
389 between academic outcomes and GRE scores. Even when accounting for the variability in these  
390 estimates (i.e., the width of the 95% CI) we see that the data support only very minor  
391 associations, if any. This can be visualized by looking at the confidence bands for the regression  
392 lines. For example, when modeling the number of first author publications as a function of  
393 quantitative GRE score, we found the rate ratio (slope) was 1.004 (95% CI 0.994 to 1.014). This  
394 implies that the average change in the number of first author publications is nearly zero even for  
395 a large shift in the GRE quantitative percentile. However, the data support changes of  
396 approximately [-1 to +1] publication. While not exactly zero, these limited data clearly support  
397 the hypothesis that there is only a very minor relationship, if any, between publication and GRE  
398 scores. In fact, for verbal scores we observed a very small negative relationship indicating (not  
399 statistically significant from zero) that there is essentially no association in these data. Similar  
400 findings can be observed for the other outcome metrics presented here, including first author

401 papers, fellowships obtained, time to degree, and faculty evaluations at exit. Importantly, we did  
402 observe a statistically significant relationship in the opposite direction with GRE and ranking  
403 (better ranked individuals tended to have poorer GRE scores). So while the overall sample size is  
404 small, there appears to be enough precision or power in these data to detect strong associations if  
405 they existed.

406

407 We have evaluated verbal and quantitative GRE scores separately in this study, but in  
408 actuality a student's application contains both scores. Perhaps a very low score in one domain  
409 (Q or V) may be offset by a high score in the other. In fact, most of the students in this cohort  
410 had two reasonably comparable Q and V scores. Of the 25 students, only four had a percentile  
411 spread between their two scores of greater than 30. In other words, they were generally either  
412 poor test takers or strong ones. Furthermore, only six of the students who completed PhDs had  
413 both GRE-Q and GRE-V scores above the 50<sup>th</sup> percentile, making it questionable whether the  
414 other 19 would have gained admittance to a graduate program that adhered to higher  
415 expectations for GRE performance. Five of the 25 students had neither GRE-Q or GRE-V  
416 scores above the 30<sup>th</sup> percentile. We think it unlikely that they would be offered admission to  
417 most graduate programs at the time or even to many programs today. Yet, among this group of  
418 five is the student who garnered the best (lowest score) faculty evaluation. These outcomes  
419 underscore the benefit of giving letters of recommendation, personal statements, and interviews  
420 far more weight than GRE scores in making admissions decisions. Our GRE-tolerant approach  
421 for increasing the number of students from historically underrepresented groups completing  
422 PhDs has been highly successful.

423

424           The relationship between objective test scores and performance has been a subject of  
425   debate for many years. Uncertainty surrounding their predictive ability must be weighed against  
426   the cost imposed on applicants to take the test, and the advantages available to a subset of  
427   applicants who can prepare extensively ahead of time and/or take the test multiple times to  
428   obtain the desired high scores. However, the outcomes of the cohort presented here indicate that  
429   non-quantitative measures (letters of recommendation, personal statements, interviews) are  
430   capable of selecting successful PhD candidates, even when those candidates have extremely low  
431   GRE scores. Subjective measures have their own drawbacks, and we sought to minimize these  
432   by having multiple, experienced readers of graduate student applications. We attempted to  
433   mediate individual biases by including multiple diverse viewpoints of each student's potential in  
434   reaching a decision to offer admission. Admittedly, this process is time consuming, but the  
435   decision of who to train as the next generation of PhD scientists is also arguably one of the most  
436   important we make.

437

438           The "GRExit" movement is growing, and for those biomedical programs that remain  
439   undecided, the data here may be helpful in arriving at a decision on whether or not to continue to  
440   require GRE scores for admission. However these decisions turn out, we assert that our GRE-  
441   tolerant approach (no score too low) undoubtedly opened doors of opportunity for PhD training  
442   at Vanderbilt that may have otherwise remained closed for historically underrepresented students  
443   with very low GRE scores. The increased diversity they bring to the community of PhD  
444   biomedical scientists will be a benefit for decades to come.

445

446

## 447 Acknowledgments

448 We thank Richard Hoover, Associate Dean of the Graduate School, for his unwavering  
449 support of the IMSD program and for working with the IMSD admissions committee to  
450 matriculate IMSD students into the Graduate School. We also thank Cathleen Williams, IMSD  
451 program manager during 2007-11, for her invaluable help and support. Finally we thank the  
452 supportive faculty mentors and the talented IMSD students from this period for their hard work  
453 and achievements that have been reported in this study. This study was supported by NIH  
454 R25GM062459 to RC and LS.

455

## 456 References

- 457 1. Moneta-Koehler L, Brown AM, Petrie KA, Evans BJ, Chalkley R. The Limitations of the  
458 GRE in Predicting Success in Biomedical Graduate School. Amaral LAN, ed. *PLoS*  
459 *ONE*. 2017;12(1):e0166742. doi:10.1371/journal.pone.0166742.
- 460 2. Hall JD, O'Connell AB, Cook JG. Predictors of Student Productivity in Biomedical  
461 Graduate School Applications. Heneberg P, ed. *PLoS ONE*. 2017;12(1):e0169121.  
462 doi:10.1371/journal.pone.0169121.
- 463 3. Hall JD. Bio/Biomedical Graduate Programs Dropping GRE Requirement. @jdhallphd  
464 Twitter Communication. November 9, 2017. Retrieved  
465 from [https://docs.google.com/spreadsheets/d/1MYcxZMhf97H5Uxr2Y7XndHn6eEC5oO](https://docs.google.com/spreadsheets/d/1MYcxZMhf97H5Uxr2Y7XndHn6eEC5oO8XWQi2PU5jLxQ/edit#gid=0)  
466 [8XWQi2PU5jLxQ/edit#gid=0](https://docs.google.com/spreadsheets/d/1MYcxZMhf97H5Uxr2Y7XndHn6eEC5oO8XWQi2PU5jLxQ/edit#gid=0)
- 467 4. <https://medschool.vanderbilt.edu/bret/igp-qcb-admissions-and-outcomes-data>

- 468 5. Miller C, Stassun K. A test that fails. *Nature*. Nature Publishing Group; 2014. June  
469 12;510(7504):303–4.
- 470 6. Posselt J. *Inside Graduate Admission: Merit, Diversity, and Faculty Gatekeeping*. 2016.
- 471 7. Weiner OD. How should we be selecting our graduate students? Drubin DG,  
472 ed. *Molecular Biology of the Cell*. 2014;25(4):429-430. doi:10.1091/mbc.E13-11-0646.
- 473 8. Bell SM, Blumstein J, Brose K, et al. Defining success in graduate school. *Molecular*  
474 *Biology of the Cell*. 2014;25(13):1942-1944. doi:10.1091/mbc.E14-03-0793.
- 475 9. Jaschick S. An Unlikely Campaign to Move Beyond GRE Scores. *Inside Higher Ed*.  
476 2016. Available: [https://www.insidehighered.com/news/2016/06/06/ets-plans-encourage-](https://www.insidehighered.com/news/2016/06/06/ets-plans-encourage-graduate-departments-de-emphasize-gre)  
477 [graduate-departments-de-emphasize-gre](https://www.insidehighered.com/news/2016/06/06/ets-plans-encourage-graduate-departments-de-emphasize-gre)
- 478 10. Benderly, BL. GREs don't predict grad school success. What does? Available:  
479 [http://www.sciencemag.org/careers/2017/06/gres-dont-predict-grad-school-success-what-](http://www.sciencemag.org/careers/2017/06/gres-dont-predict-grad-school-success-what-does)  
480 [does](http://www.sciencemag.org/careers/2017/06/gres-dont-predict-grad-school-success-what-does)
- 481 11. Park H-Y, Berkowitz O, Symes K, Dasgupta S. The art and science of selecting graduate  
482 students in the biomedical sciences: Performance in doctoral study of the foundational  
483 sciences. Prunuske A, ed. *PLoS ONE*. 2018;13(4):e0193901.  
484 doi:10.1371/journal.pone.0193901.

485

486

## 487 Supporting information

488 **Table S1.** (corresponds to Fig 3. Associations between quantitative and verbal GRE scores and  
489 total number of publications)

490 Results from Poisson regression models looking at the association between GRE-Quantitative and number of  
491 publications (Table 1a) and GRE-Verbal and number of publications (Table 1b). The columns show the estimated  
492 rate ratios, model robust standard errors, 95% confidence intervals, and p-values.

493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
  
520  
  
521  
  
522  
  
523  
  
524

**Table S2.** (corresponds to Fig 4. Associations between quantitative and verbal GRE scores and total number of first author publications)

Results from Poisson regression models looking at the association between GRE-Quantitative and number of first author publications (Table 2a) and GRE-Verbal and number of first author publications (Table 2b). The columns show the estimated rate ratios, model robust standard errors, 95% confidence intervals, and p-values.

**Table S3.** (corresponds to Fig 5. Associations between quantitative and verbal GRE scores and months to degree)

Results from Poisson regression models looking at the association between GRE-Quantitative and number of months to degree (Table 3a) and GRE-Verbal and months to degree (Table 3b). The columns show the estimated rate ratios, model robust standard errors, 95% confidence intervals, and p-values.

**Table S4.** (corresponds to Fig 6. Boxplots of GRE scores stratified by whether or not the students received a fellowship)

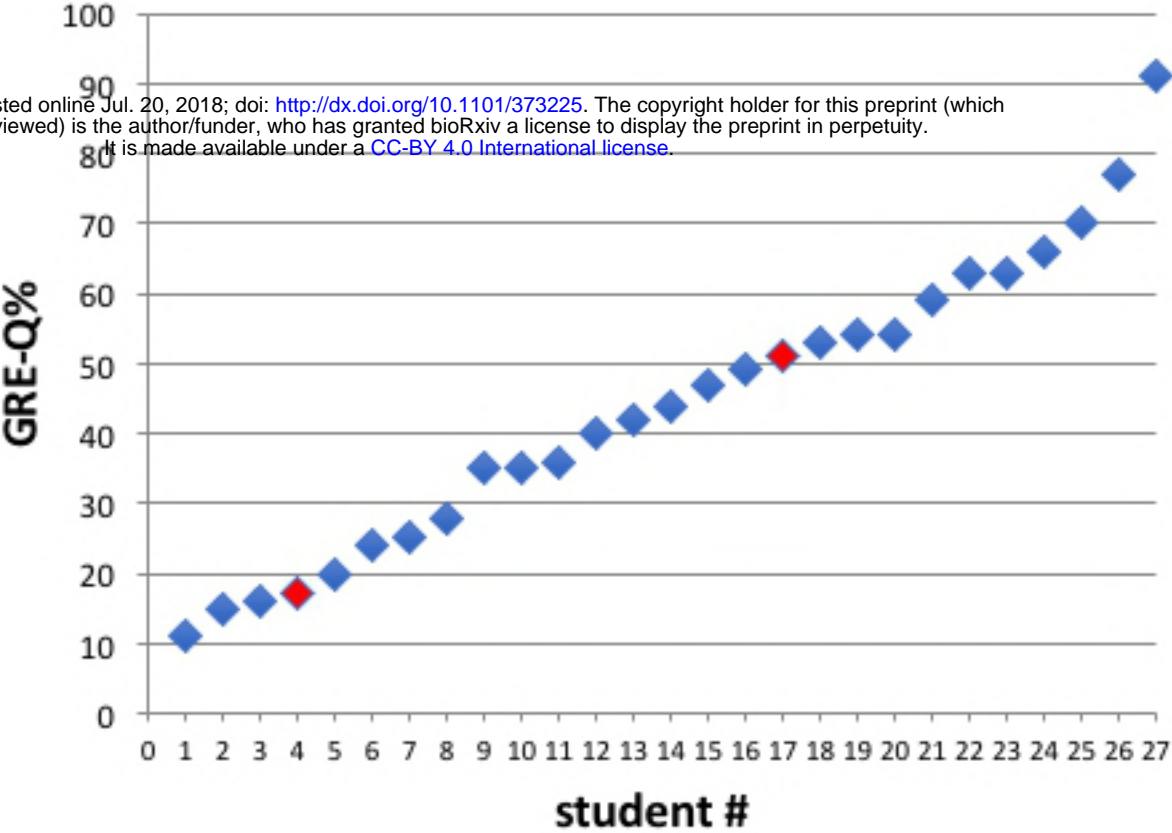
Results from logistic regression models looking at the association between GRE-Quantitative and receipt of fellowship (Table 4a) and GRE-Verbal and receipt of fellowship (Table 4b). The columns show the estimated odds ratios, model robust standard errors, 95% confidence intervals, and p-values.

**Table S5.** (corresponds to Fig 8. Associations between GRE scores and faculty ranking)

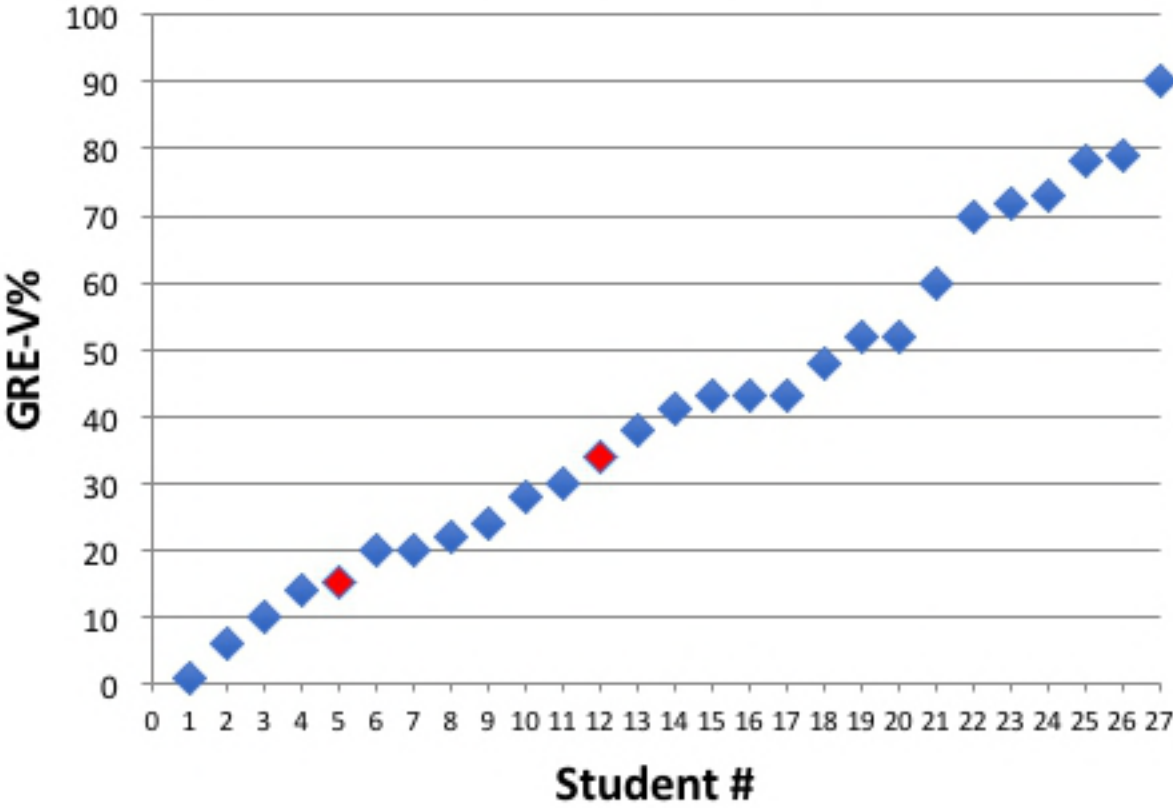
Results from Poisson regression models looking at the association between GRE-Quantitative and faculty ranking (Table 5a) and GRE-Verbal and faculty ranking (Table 5b). The columns show the estimated rate ratios, model robust standard errors, 95% confidence intervals, and p-values.

# GRE-Q% PhD+Masters

bioRxiv preprint first posted online Jul. 20, 2018; doi: <http://dx.doi.org/10.1101/373225>. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

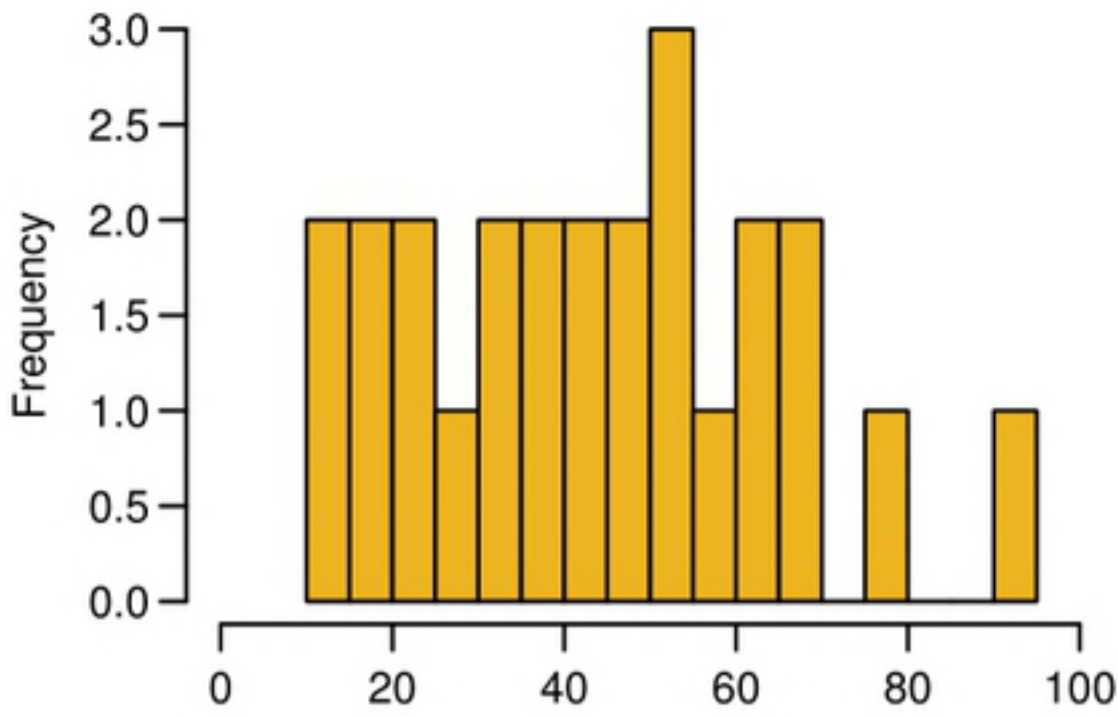


# GRE-V% PhD+Masters

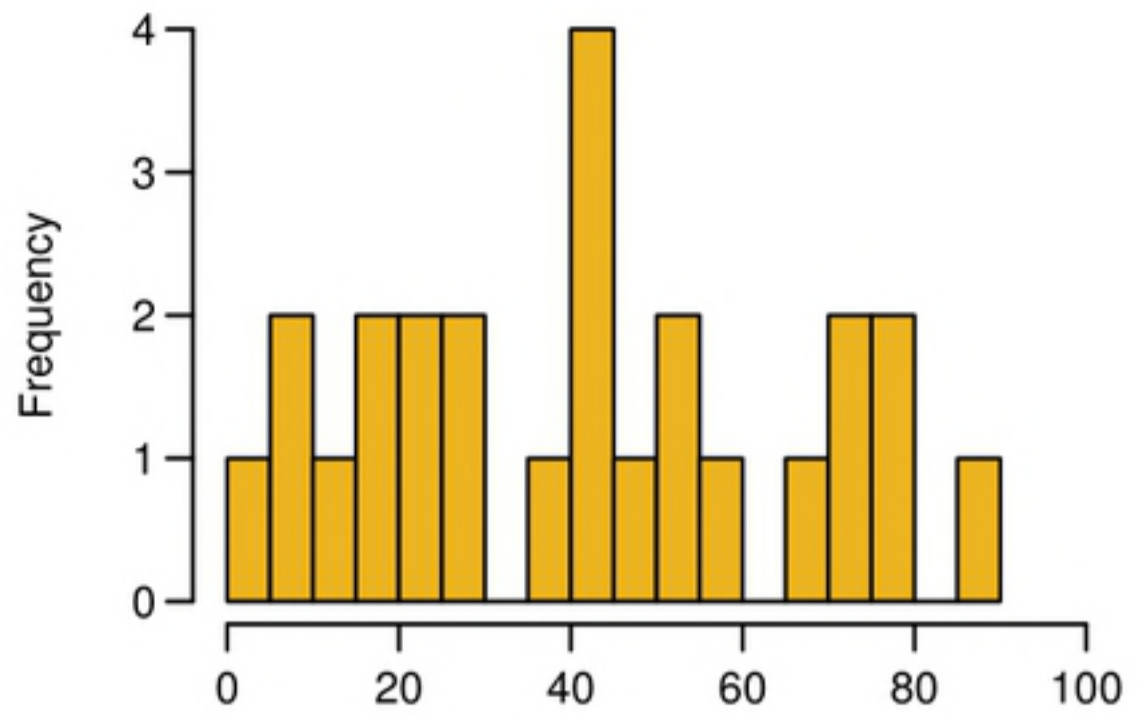




### GRE-Quantitative

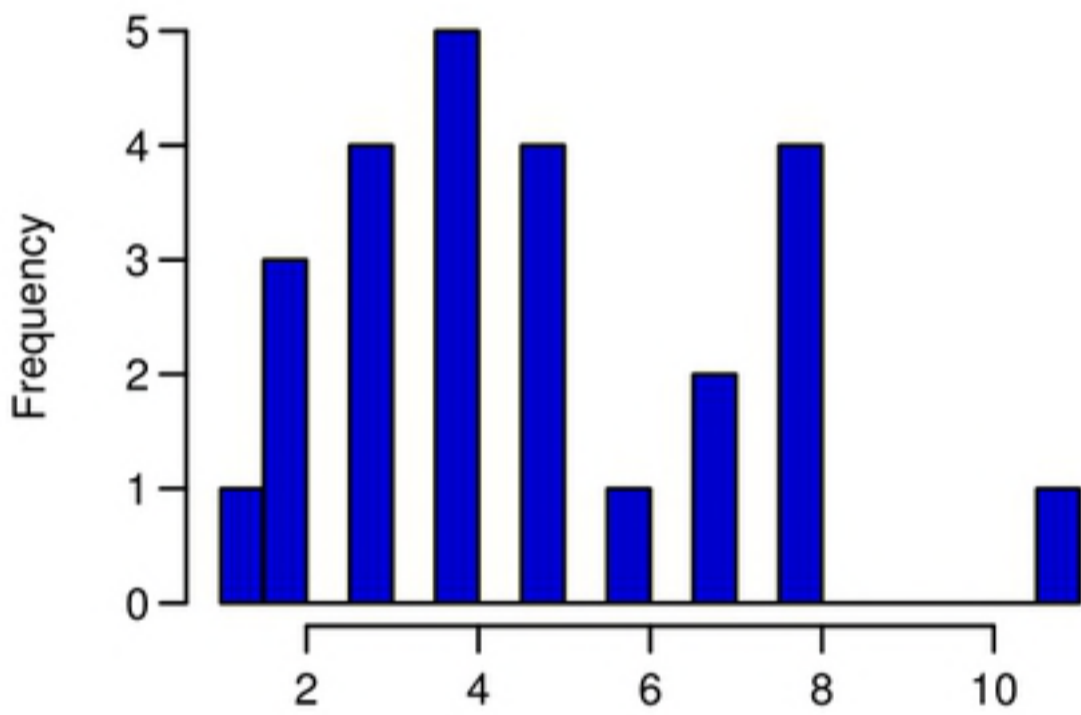


### GRE-Verbal

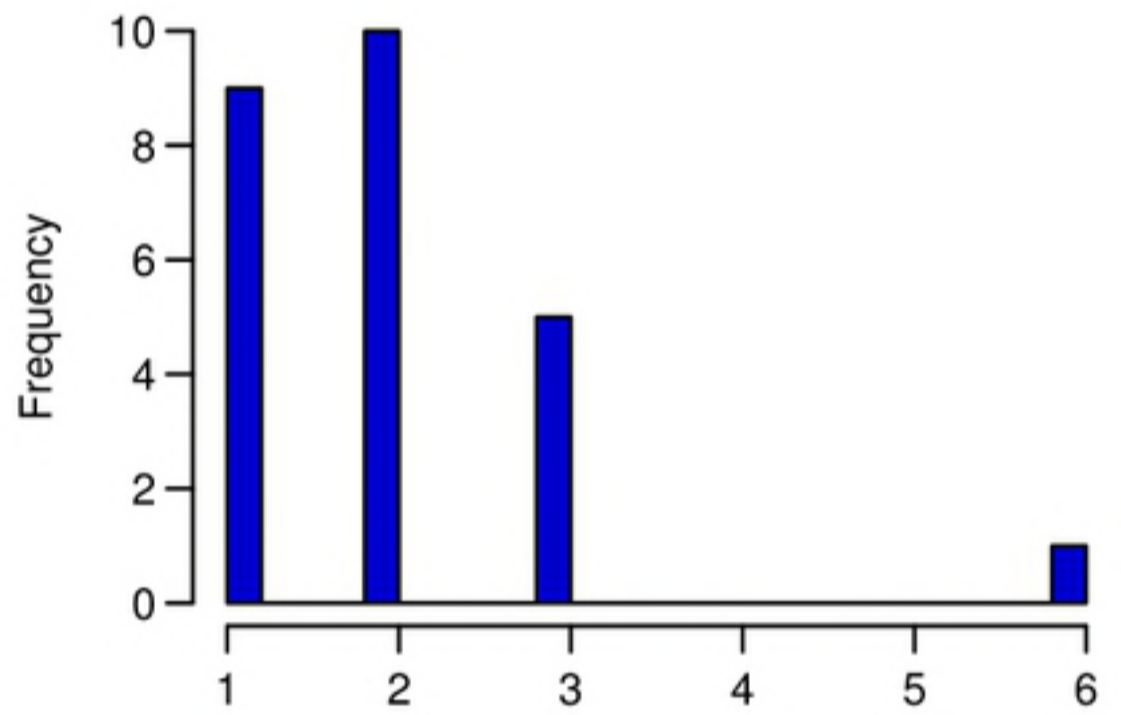


bioRxiv preprint first posted online Jul. 20, 2018; doi: <http://dx.doi.org/10.1101/373225>. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

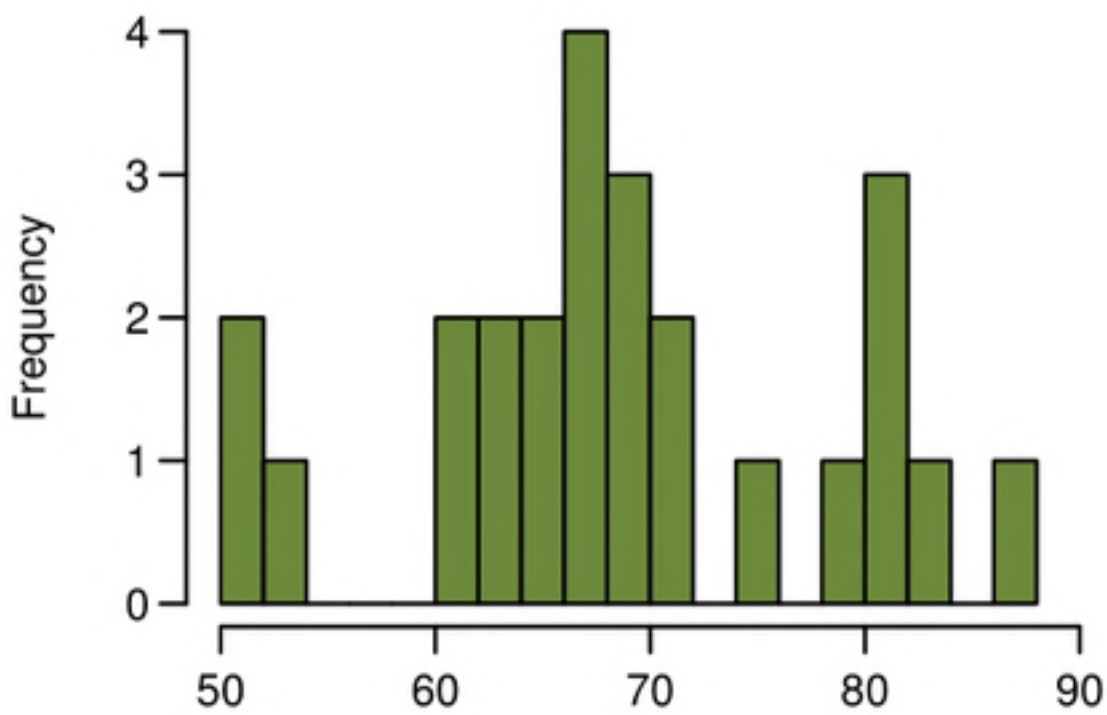
### Number of publications



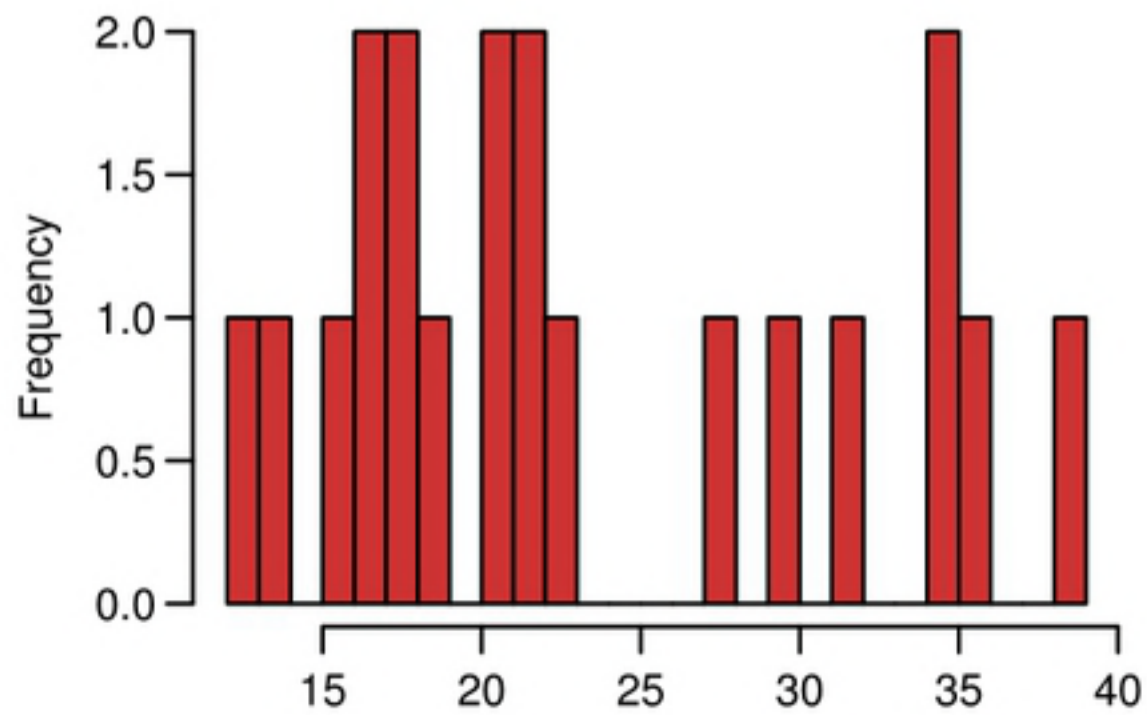
### First author publications

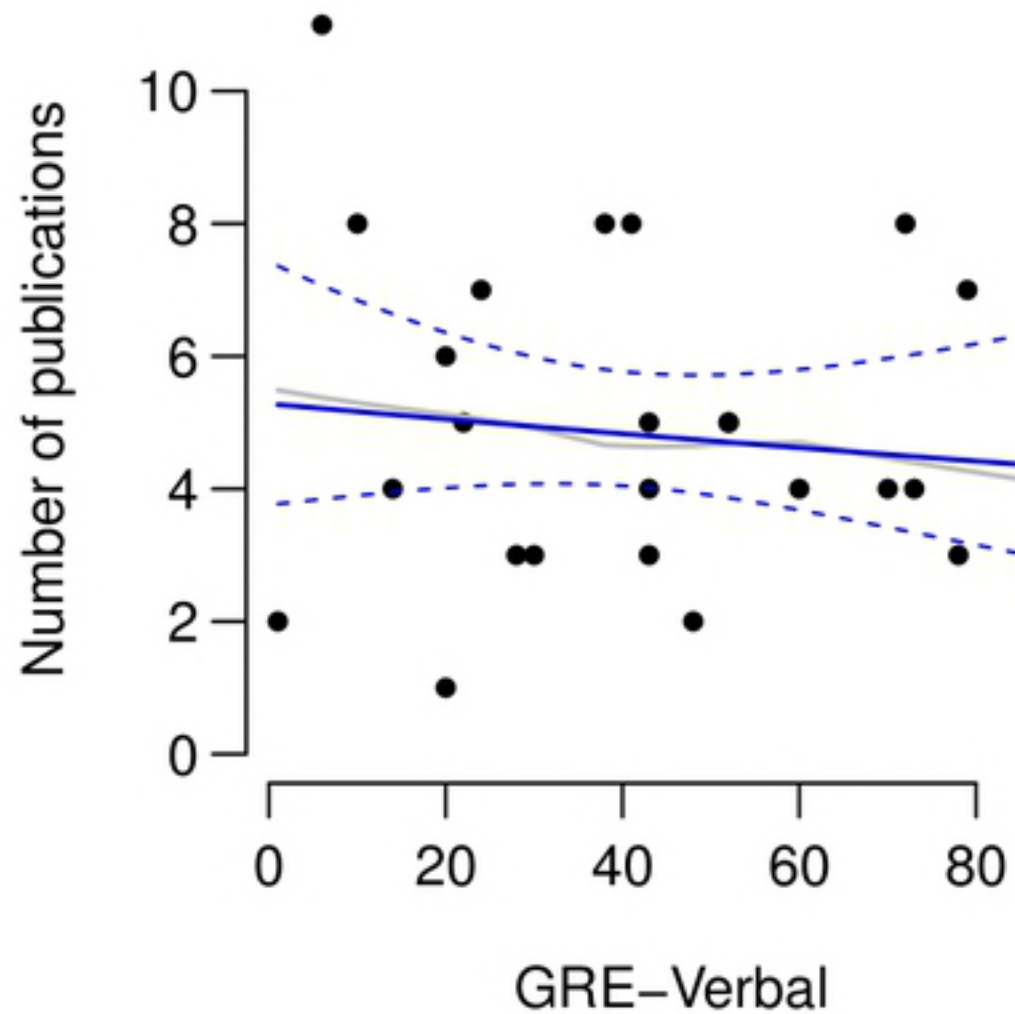
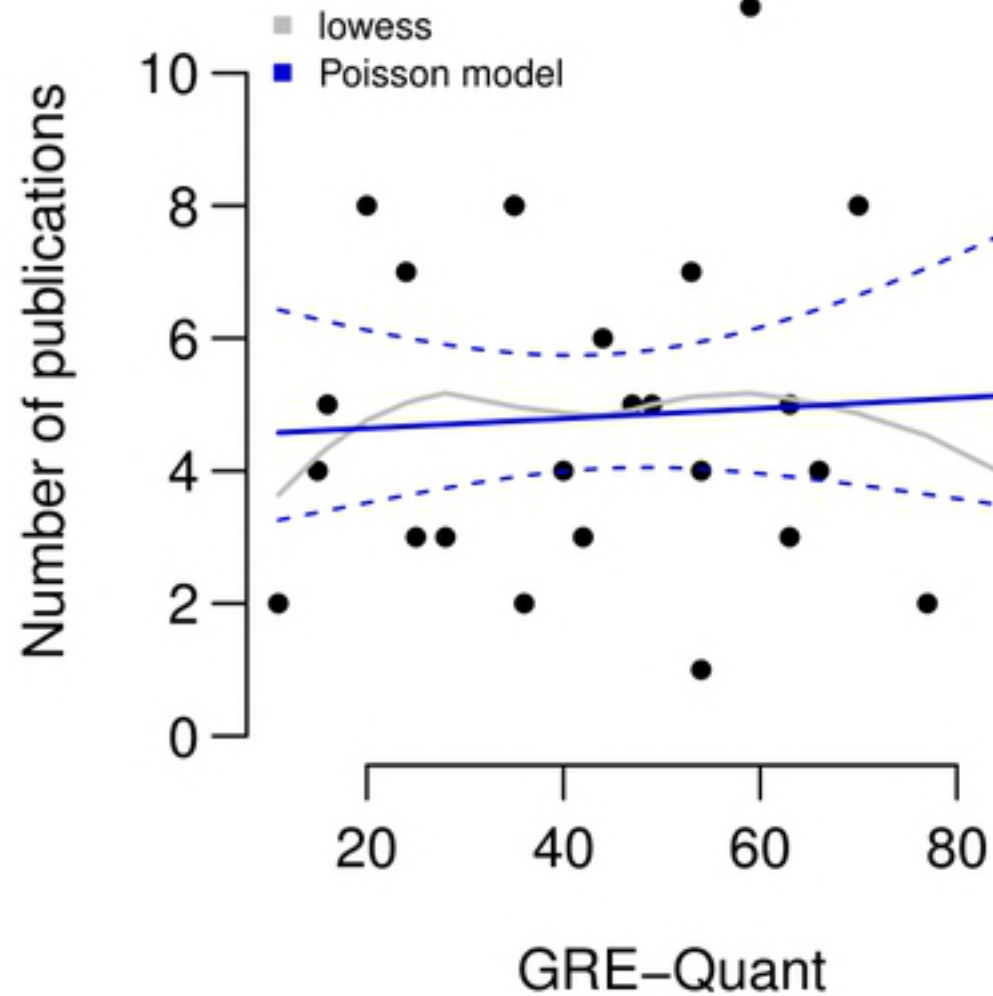


### Months to degree

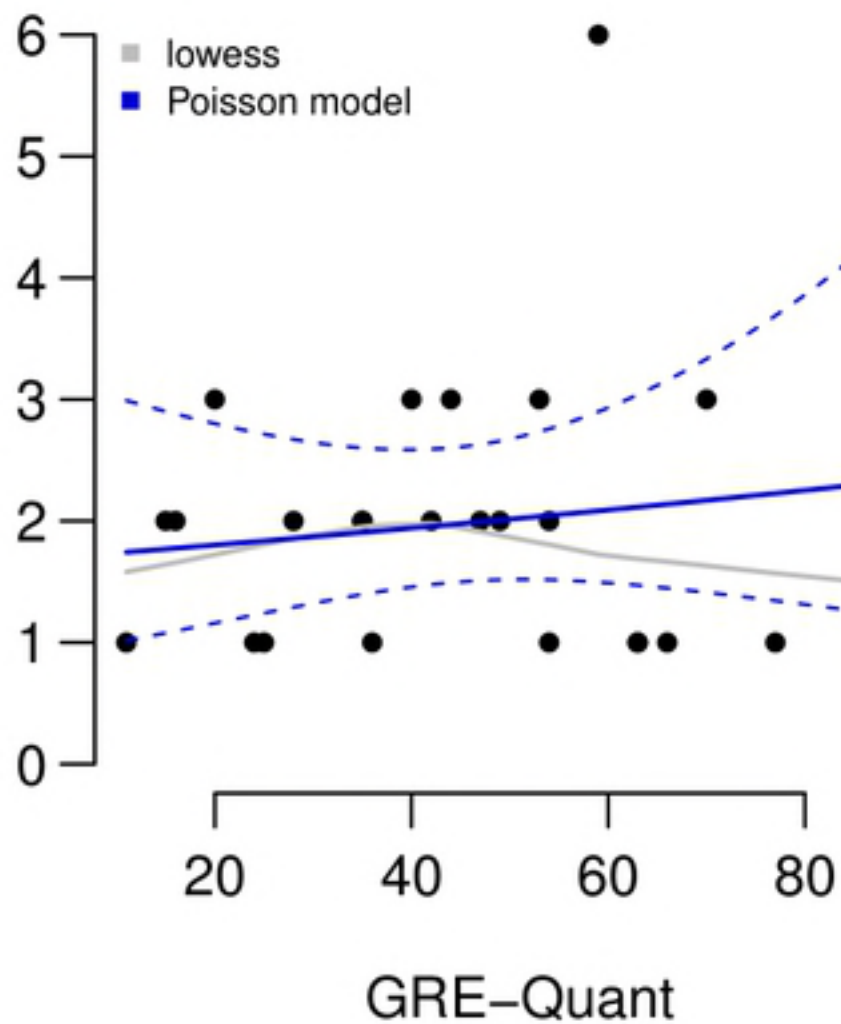


### Faculty ranking

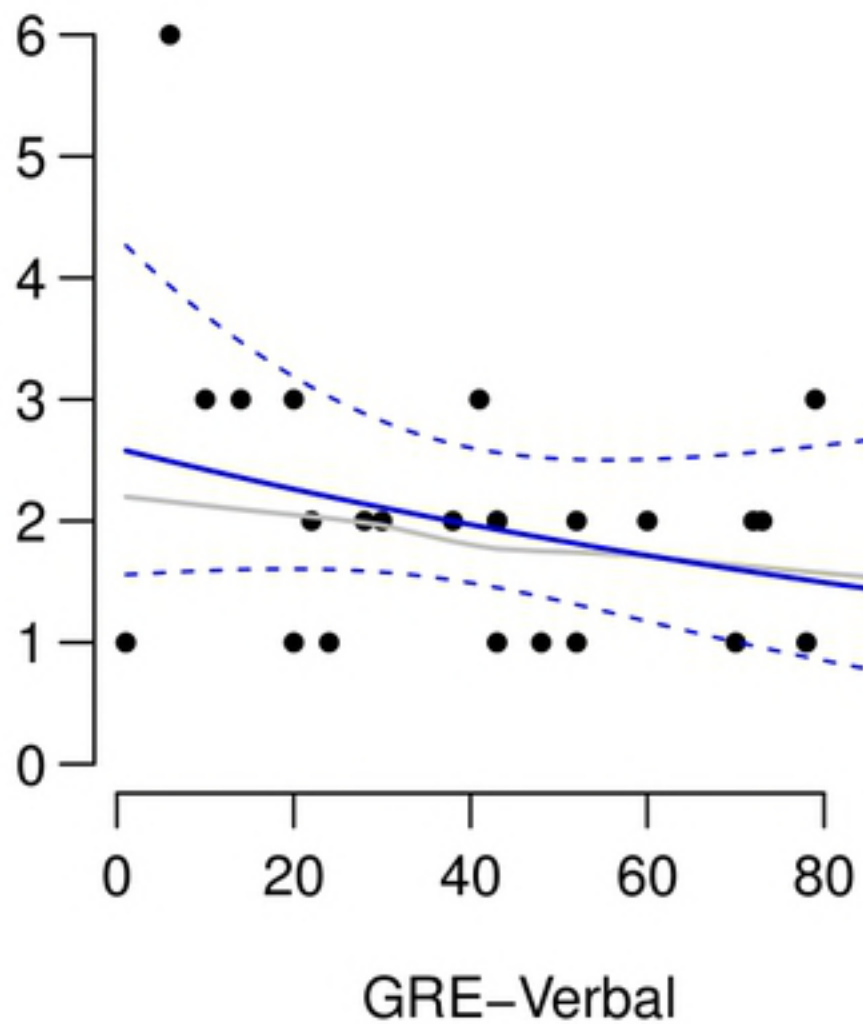


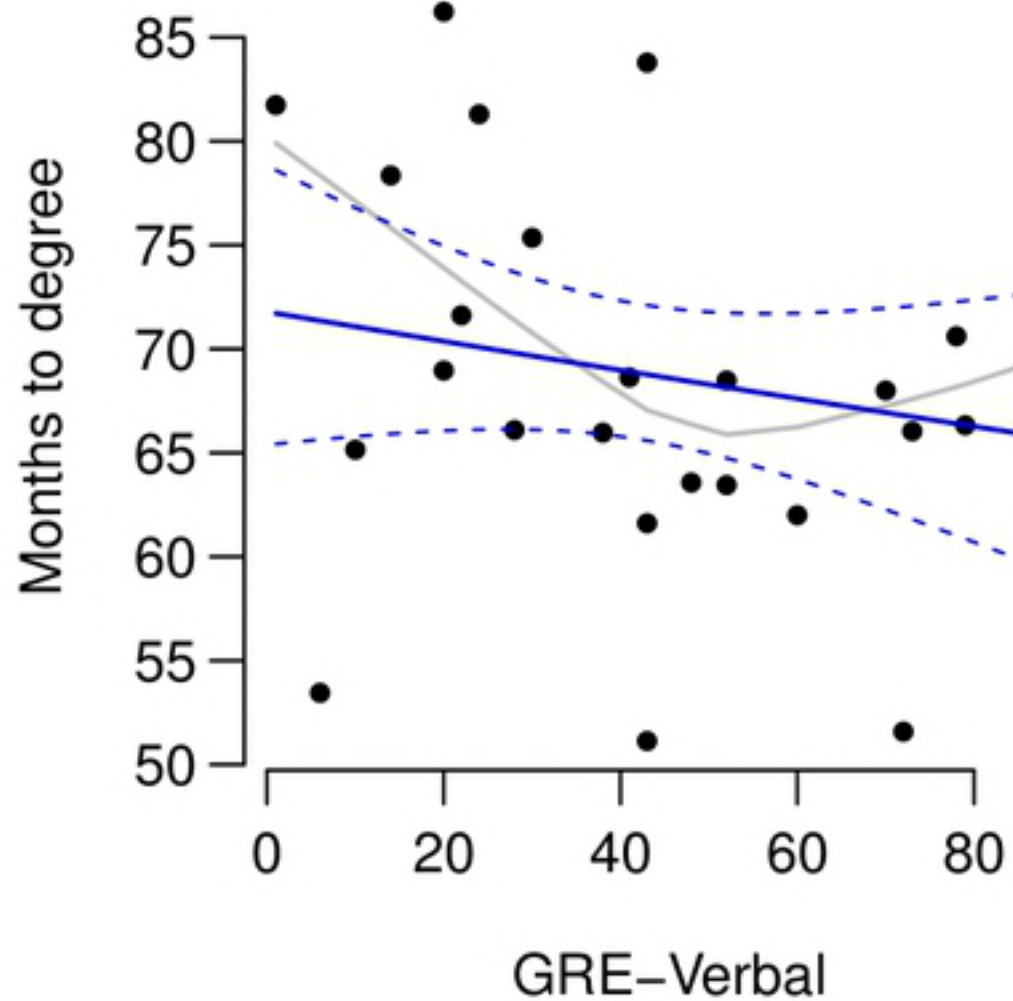
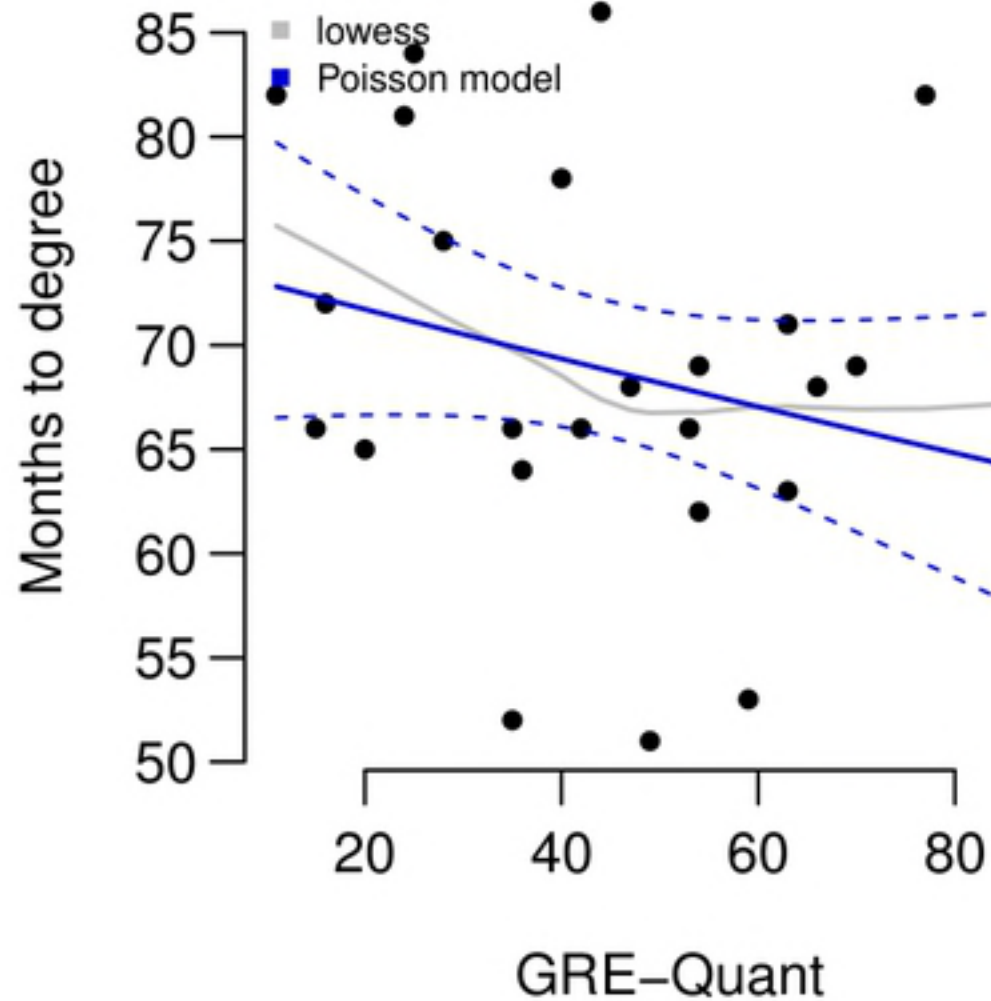


First author publications

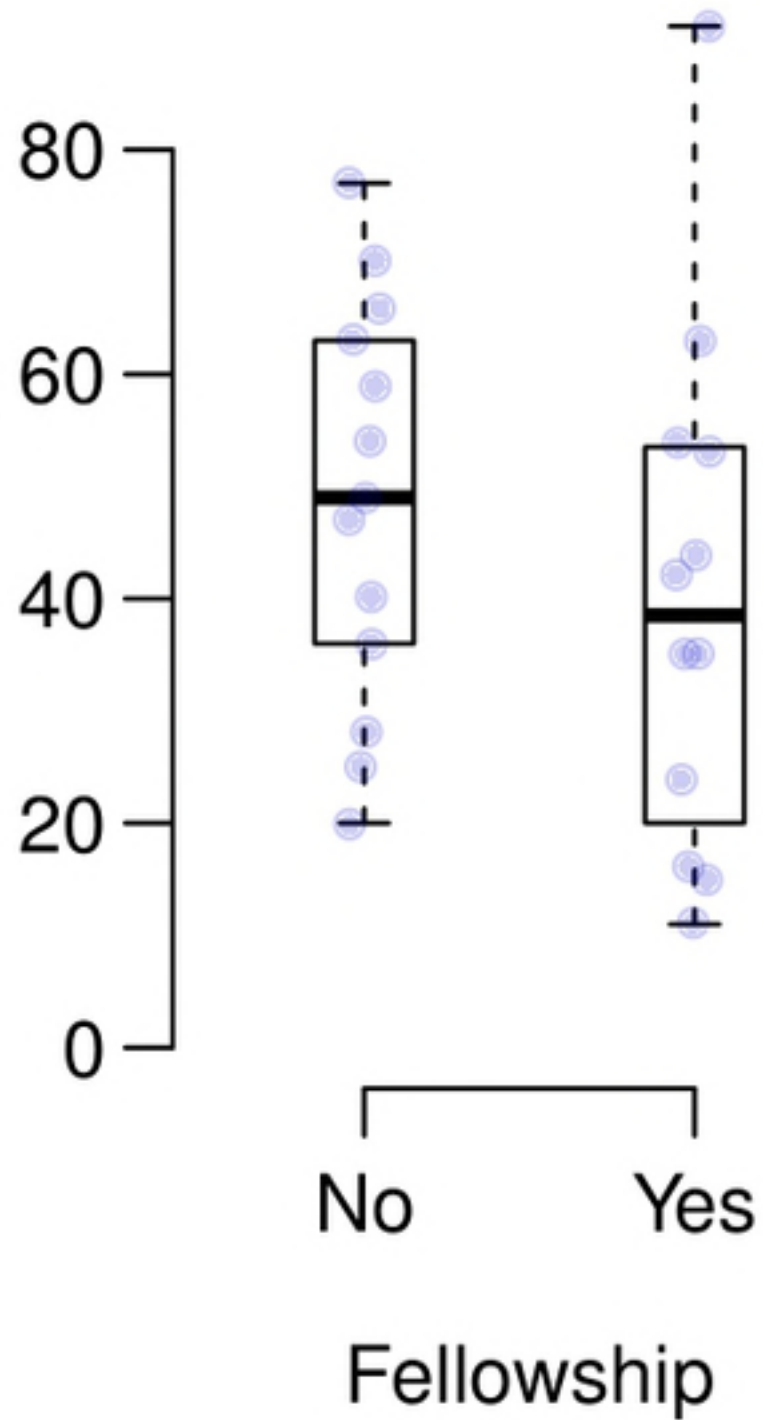


First author publications





**GRE-Quant**



**GRE-Verbal**

