**Supplemental Digital Content**

1. **Initial recruitment email**

Dear Dr. [potential participant’s last name],

My name is Laura Corlin and I am an Assistant Professor at Tufts University School of Medicine. I am writing to invite you to participate in a research study on the academic training of researchers publishing peer-reviewed epidemiological papers. The information we collect from this survey will be used in a publication and will inform the development of resources to improve training in epidemiology and biostatistics for researchers. You are eligible to participate in this study because you are an author on [insert published paper name] published in [year] in [journal]. If you choose to participate, you will be asked to complete a survey [link] that should take no more than 5-10 minutes of your time.

If you are interested in the study, please click here [link]. If you have any questions, please feel free to reach out to me by email (laura.corlin@tufts.edu) or phone (617-636-0463).

Additionally, the Tufts University Data Protection Notice is below for you to read [participants were given the appropriate form].

Thank you very much for your help!

Sincerely,

Laura Corlin, PhD

[Title and contact information given]

1. **Follow-up recruitment email**

Dear Dr. [potential participant’s last name],

I emailed you last week to invite you to participate in a research study on the academic training of researchers publishing peer-reviewed epidemiological papers. If you have already completed the short (5-10 minute) survey, thank you and please ignore this email. If not, we hope that you will consider participating.

You are being contacted because you are an author on [insert published paper name] published in [year] in [journal]. If you choose to participate, you will be asked to complete a short survey [link].

If you are interested in the study, please click here [*link*]. If you have any questions, please feel free to reach out to me by email (laura.corlin@tufts.edu) or phone (617-636-0463).

Additionally, the Tufts University Data Protection Notice is below for you to read [participants were given the appropriate form].

Thank you!

Sincerely,

Laura Corlin, PhD

[Title and contact information given]

1. **Survey question text**
2. What is your gender? [participants could select one option]
* Female
* Male
* Non-Binary
* Other (Please describe)
* Prefer not to answer
1. Do you identify as transgender? [participants could select one option]
* Yes
* No
* Prefer not to answer
1. What is your year of birth (yyyy)? [participants could type a four-digit answer]
2. As of the time your paper was published, what is the most advanced degree you had been awarded above a bachelor’s degree? [participants could select one option]
* Master's
* PhD
* MD
* PhD and MD
* Other doctoral level degree
* Other
* None of the above
1. As of the time your paper was published, in what year had your most recent degree been awarded? [participants could type a four-digit answer]
2. As of the time your paper was published, in what field(s) had you obtained advanced degree(s)? (you can choose more than one)
* Epidemiology
* Biostatistics
* Medicine/clinical training
* Other public health field (please specify) [participants could type a response]
* Other [participants could type a response]
1. As of the time your paper was published, in what country (or countries) had you received your educational training? [participants could type a response]
2. As of the time your paper was published, had you ever received formal training in epidemiology? (Select all that apply)
* Yes, at least a full semester of an undergraduate-level course
* Yes, at least a full semester of a graduate-level course
* Yes, in a workshop
* Yes, in an online class
* Yes, in another place
* No
* Unsure
* Prefer not to answer
1. If answered no, unsure, or prefer not to answer to question 8: As of the time your paper was published, would you have benefited from formal training (at least a full semester course) in epidemiology? [participants could select one option]
* Yes
* No
* Unsure
* Prefer not to answer
1. If answered yes to question 8: As of the time your paper was published, in what year had you most recently obtained formal epidemiology training? [participants could type a four-digit answer]
2. As of the time your paper was published, had you ever received formal training in biostatistics? (Select all that apply)
* Yes, at least a full semester of an undergraduate-level course
* Yes, at least a full semester of a graduate-level course
* Yes, in a workshop
* Yes, in an online class
* Yes, in another place
* No
* Unsure
* Prefer not to answer
1. If answered no, unsure, or prefer not to answer to question 11: As of the time your paper was published, would you have benefited from formal training (at least a full semester course) in biostatistics? [participants could select one option]
* Yes
* No
* Unsure
* Prefer not to answer
1. If answered yes to question 11: As of the time your paper was published, in what year had you most recently obtained formal biostatistics training? [participants could type a four-digit answer]
2. As of the time your paper was published, how confident were you in your personal ability to appropriately:
	1. Apply epidemiological concepts relevant to your paper [participants could select not at all, not very, somewhat, very, or extremely]
	2. Apply biostatistical concepts relevant to your paper [participants could select not at all, not very, somewhat, very, or extremely]

In this section, we will be asking about your co-authors on the paper. Exclude your own training when answering these questions. For this section, formal training is defined as at least two full semester courses in the subject.

1. As of the time your paper was published, were any of your co-authors formally trained in epidemiology? [participants could select one option]
* Yes
* No
* Unsure
1. If answered yes to question 15: In what stage of the project did the epidemiologist(s) participate? (Select all that apply)
* Design
* Analysis
* Interpretation
* Reporting
* Other [participants could type a response]
1. As of the time when your paper was published, were any of your coauthors formally trained in biostatistics? [participants could select one option]
* Yes
* No
* Unsure
1. If answered yes to question 17: In what stage of the project did the biostatistician(s) participate? (Select all that apply)
* Design
* Analysis
* Interpretation
* Reporting
* Other [participants could type a response]
1. As of the time your paper was published, did any of your co-authors have a clinical degree? [participants could select one option]
* Yes
* No
* Unsure
1. If answered yes to question 19: In what stage of the project did the clinician(s) participate? (Select all that apply)
* Design
* Analysis
* Interpretation
* Reporting
* Other
1. If answered yes to question 19: As of the time your paper was published, did any of your clinician coauthors have formal training in epidemiology and/or biostatistics? [participants could select one option]
* Yes
* No
* Unsure
1. Prior to submitting your paper for publication in the journal it was accepted in, did anyone formally trained in epidemiology and/or biostatistics review your paper? [participants could select one option]
* Yes
* No
* Unsure
1. As of the time your paper was published, how confident were you in your co-authors’ ability to appropriately:
2. Apply epidemiological concepts relevant to your paper [participants could select not at all, not very, somewhat, very, or extremely]
3. Apply biostatistical concepts relevant to your paper [participants could select not at all, not very, somewhat, very, or extremely]
4. As a reminder, we are trying to learn about the training of authors publishing epidemiologic research. Do you have any final thoughts you would like to share with us on this topic? [participants could type a response]

**Table S1. P-values for the associations between independent variables and interdisciplinary training1 and team composition2**

|  |  |  |
| --- | --- | --- |
| **Independent variable** | **Interdisciplinary training** | **Interdisciplinary team composition** |
| General clinical versus epidemiology journal (referent) | 0.010 | 0.233 |
| Specialty clinical versus epidemiology journal (referent) | 0.052 | 0.227 |
| General or specialty clinical versus epidemiology journal (referent) | 0.011 | 0.157 |
| Article year – 2000 versus 2020 (referent) | 0.596 | 0.768 |
| Article year – 2010 versus 2020 (referent) | 0.342 | 0.471 |
| Article year – 2010 versus 2000 (referent) | 0.580 | 0.584 |
| Gender of corresponding author (female/male) | 0.155 | 0.993 |
| Age of corresponding author at time of publication | 0.351 | 0.203 |
| Corresponding author’s epidemiology training (y/n) | NA | 0.001 |
| Corresponding author’s recency of epidemiology training at time of publication (≥/<5 years) | 0.684 | 0.314 |
| Corresponding author’s biostatistics training (y/n) | NA | 0.093 |
| Corresponding author’s recency of biostatistics training at time of publication (≥/<5 years) | 0.581 | 0.246 |

1Interdisciplinary training = a clinical degree and personal epidemiology training and personal biostatistics training
2Interdisciplinary team composition = among all co-authors (including the corresponding author), at least one person has clinical training, at least one person has epidemiology training, and at least one person has biostatistics training

**Table S2. Open-ended responses to the question: “As a reminder, we are trying to learn about the training of authors publishing epidemiologic research. Do you have any final thoughts you would like to share with us on this topic?”**

|  |  |
| --- | --- |
| **Theme** | **Response**  |
| Respondent's training  | this paper was published while I was in process of defending my PhD dissertation in Epidemiology, and my doctorate degree was awarded less than 2 months after this paper came out. |
| Work conducted as a PhD student, supervised by world leader in longitudinal data analysis [redacted name and degree] |
| My training is not in biostats/epi but in another field where the math rigor, stats, and modeling is at a more rigorous and technical level than what is formally taught in public health or medical schools |
| After about 5 years from my PhD I worked for a year with [redacted name]. I think that was probably the best training I could have had! I also subsequently worked with [redacted name], so again learnt a lot |
| I basically trained myself in clinical epidemiology and biostatistics. |
| I think the emphasis on formal training in this survey is problematic. I have systematically avoided courses in biostatistics and epidemiology and instead have learned through close collaborations with researchers who have such training. For me that is a more effective way of learning. I think I'm not the only one. |
| My understanding of epidemiology and biostatistics is largely from working in a team of doctorate level epidemiologists and biostatisticians over approximately 30 years. |
| Since that paper was published over 20 years ago, I have received much more research training but still lag in epidemiology skills. |
| We all were quite well trained in clinical epidemiology in the beginning in the [redacted year and university] and I have always found it extremely helpful to have a sound epidemiological background although having a clinical background. thereafter I personally had some more intense course on Regression analysis which has also been extremely helpful |
| This was a fairly straightforward pooled analysis. In retrospect, I would have used a random effects meta-analysis statistical approach. But either way, the conclusion would have been the same because of the extremely low incidence of the outcome of interest. |
| Co-authors' training | We formed a multi-author multidisciplinary team that specifically included a highly qualified epidemiologist and individuals with extensive experience in survey design, as well as public health and clinical experience. I don't think it is possible for any one individual to hold all the expertise needed for this sort of project. |
| This paper happened to have extensive Epi training and mix of related training: [redacted list of training of individual authors] |
| the paper was part of a PhD training and became a chapter in a PhD thesis. All co-authors were members of the department of Epidemiology and Biostatistics |
| My doctorate is in epidemiology, one of the co-authors is a doctoral candidate and we're both trained in biostats as part of the epi PhD. I would guess that you're trying to understand whether papers with epi and analytical content consult with trained epidemiologists and biostatisticians. Given that, when the affiliation of two of the authors is a department of epidemiology in a school of public health then there is an increased likelihood that those authors will have training in epi |
| This paper didn't have a co-author who was primarily an epidemiologist or a biostatistician, but the senior author was a clinician who had MPH training in epi and biostats, as I did. |
| You have picked an interesting example where, unusually for our group, there was no formally trained statistician on the team. The three clinicians [redacted specialties] were all very experienced researchers and two of them had done additional short courses in epidemiology and in statistics but, as far as I can recall, did not have postgraduate qualifications in either. I think this style of collaboration at an individual level is now much rarer - projects like this would now expect to have formal statisticians and information scientist on board. |
| my coauthors had training in health services research and pharmaceutical health outcomes research. I think training in these fields is also very applicable for applied epidemiology/policy analysis and you may be under-representing this expertise if just focusing on formal epidemiology and biostatistical training. |
| That paper represented an active collaboration between scientists working in a government agency and academic colleagues from academia. It was teamwork. Authors in public health need to learn how to work in multi-disciplinary teams. |
| I am fortunate to be part of a very diverse and highly experienced team |
| Challenges in reporting on training | Epidemiology overlaps a \*lot\* with quite a few other social sciences!!!!!!! |
| You talk about Biostatistics but I studied Statistics. Many older statisticians like me studied at a time when the term hadn't been invented. Indeed, the term is particularly American - I am now Prof of [redacted field] in [redacted non-US country] and would never call myself a biostatistician. Your survey might, thus, misrepresent the prevalence of (bio)statisticians |
| You need to \*DEFINE\* biostatistics, and possible even EPIDEMIOLOGY before starting to ask question about the topics. For example, a Master's degree in mathematics focusing on mathematical statistics, is that biostatistics or not!? Additionally, it's unclear why you probe into a paper published 10 years ago, it's very hard to remember exactly what happened in what order so long ago. |
| my coauthors had training in health services research and pharmaceutical health outcomes research. I think training in these fields is also very applicable for applied epidemiology/policy analysis and you may be under-representing this expertise if just focusing on formal epidemiology and biostatistical training. |
| Reflections on interdisciplinarity  | That paper represented an active collaboration between scientists working in a government agency and academic colleagues from academia. It was teamwork. Authors in public health need to learn how to work in multi-disciplinary teams. |
| We formed a multi-author multidisciplinary team that specifically included a highly qualified epidemiologist and individuals with extensive experience in survey design, as well as public health and clinical experience. I don't think it is possible for any one individual to hold all the expertise needed for this sort of project. |
| Training is essential. Too many people assume they know how to interpret data and as a result, the quality of many papers suffer. It’s as complex as specialty training in [redacted clinical field]! |
| I am fortunate to be part of a very diverse and highly experienced team |
| Reflections on the role of confidence | At the time, I was very young and overconfident. As it was twenty years ago and I have learned a lot since then, I would not be surprised to find errors in the paper, if I were to read it again today. |
| these questions are a bit subjective. e.g., someone with lots of confidence will check 'extremely' whereas others who are less confident will not. |
| Reflections on training programs | Most US medical school provide several curriculum lectures in Epidemiology. I[t] provides recognition, but not expertise. |
| The problem with statistics teaching in medical schools that I know is that it tends to be too theoretical and too early (you forget!) to be of real use when initiating research. |
| All students at schools of public health should be required to take a course or seminar on the importance of skepticism in the interpretation of one's research |
| I have found having formal epidemiology and biostatistics training to be extremely useful in publishing the research that I conduct. |
| I am not sure how someone who does not have epidemiology training could publish good research. This point seems obvious. Why do you need a survey to figure this out |
| We all were quite well trained in clinical epidemiology in the beginning in [redacted year and university] and I have always found it extremely helpful to have a sound epidemiological background although having a clinical background. thereafter I personally had some more intense course on Regression analysis which has also been extremely helpful |
| This was a fairly straightforward pooled analysis. In retrospect, I would have used a random effects meta-analysis statistical approach. But either way, the conclusion would have been the same because of the extremely low incidence of the outcome of interest. |
| I think the emphasis on formal training in this survey is problematic. I have systematically avoided courses in biostatistics and epidemiology and instead have learned through close collaborations with researchers who have such training. For me that is a more effective way of learning. I think I'm not the only one. |
| Other | The paper you are requesting information on was published in 2000, developed prior to that. It was a small sample trial that was uncontrolled. Not worth a lot of attention in today's world, from a statistical or design standpoint. |
| The study that you are referring to was not epidemiological in its nature. |
| this was a narrative review, not a scientific study. |
| You selected one of my papers that was published in [redacted name of epidemiology journal]. I hope your research is capturing authors of epi papers outside of the mainstream epi journals. |
| I would be more confident that I had answered the questions correctly if you had picked a more recent paper. |
| Some coauthors were put on there because of their connection to the funding source, although they did contribute. |

**Figure S1. Confidence in personal epidemiological and biostatistical abilities**



**Figure S2. Confidence in co-author epidemiological and biostatistical abilities**

