RESEARCH ARTICLE

REVISED Factors associated with producing a scientific publication during medical training: evidence from a cross-sectional study of 40 medical schools in Latin America [version 2; peer review: 2 approved]

Previously titled: Factors associated with producing a scientific publication during medical training: evidence from a cross-sectional study of 40 medical schools surveyed in Latin America

Mario J. Valladares-Garrido[1,2], Christian R. Mejia[3], Annel B. Rojas-Alvarado[4-6], Mary M. Araujo-Chumacero[7], Jhacksson S. Córdova-Agurto[5,7], Jessica Fiestas[8], Feeder J. Rojas-Vilar[9], Carlos Culquichicón[10]

[1]Vicerrectorado de Investigación, Universidad Norbert Wiener, Lima, Peru
[2]Instituto de Evaluación de Tecnologías en Salud e Investigación-IETSI, EsSalud, Lima, Peru
[3]Universidad Continental, Lima, Peru
[4]School of Medicine, Universidad Privada Antenor Orrego, Piura, Peru
[5]CI- Emerge, Center of Emerging Diseases and Climate Change, Universidad Nacional de Piura, Piura, Peru
[6]Scientific Society of Medical Students, Universidad Privada Antenor Orrego, Piura, Peru
[7]School of Health Sciences, Universidad Nacional de Piura, Piura, Peru
[8]School of Medicine, Universidad Cesar Vallejo, Piura, Peru
[9]Facultad de Medicina, Universidad Ricardo Palma, Lima, Peru
[10]Universidad Privada Norbert Wiener, Centro de Investigación Epidemiológica en Salud Global, Lima, Peru

v2 **First published:** 24 Nov 2020, **9**:1365
https://doi.org/10.12688/f1000research.26596.1
Latest published: 23 May 2022, **9**:1365
https://doi.org/10.12688/f1000research.26596.2

Abstract

Background: Scientific publication during medical training is key to promoting enduring cutting-edge knowledge. The promotion of science among medical students in Latin America is a multisectoral issue that is hampered by the lack of governmental knowledge to invest in national research, as well as by the lack of support from local universities. This study aims to determine the factors associated with the production of a scientific publication during medical training among Latin American medical students of local scientific societies.
Methods: This is a secondary data analysis of a cross-sectional study conducted in 2016 that assessed the use of information and communication technologies (ICTs) among medical students from 40 local scientific societies of medical students affiliated with FELSOCEM. Teams from each local scientific society surveyed self-reported scientific publications and explored their association with socioeconomic, academic, and research training conditions. We

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 23 May 2022		✓ view
version 1 24 Nov 2020	✓ view	? view

1. **Virgilio Efrain Failoc Rojas**, Saint Ignatius of Loyola University, Lima, Peru
2. **Megan Anakin**, University of Otago, Dunedin, New Zealand

applied nested models to identify the covariates associated with self-reported scientific publication, obtaining a parsimonious mixed-effects multilevel model grouped by medical scientific society.
Results: Of 11,587 participants, the prevalence of scientific publications increased in 36% among medical students affiliated to a Scientific Society of Medical Students [parsimonious prevalence ratio (pPR)=1.36, 95%CI=1.16–1.59], 51% among medical students with advanced English proficiency [pPR=1.51, 95%CI=1.21 – 1.87], 85% among medical students who attended a scientific writing skills course [pPR=1.85, 95%CI=1.59–2.15], 81% among medical students who use Sci-Hub [pPR=1.81, 95%CI=1.50–2.20], and 108% among medical students who have access to a pirated academic account [pPR=2.08, 95%CI=1.83–2.36].
Conclusions: Producing a scientific publication among medical students is associated with being affiliated to a scientific society of medical students, English proficiency, training in scientific writing, use of Sci-Hub, and pirated academic accounts. The results will help clinical educators and medical programs improve resources for training students in high-quality research

Keywords
Medical Education, Undergraduate, Scientific Societies, Latin America, Medical Students

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Carlos Culquichicón (carlos.culquichicon@unp.edu.pe)

Author roles: Valladares-Garrido MJ: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Mejia CR**: Conceptualization, Data Curation, Investigation, Methodology, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Rojas-Alvarado AB**: Conceptualization, Data Curation, Investigation, Methodology, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Araujo-Chumacero MM**: Conceptualization, Data Curation, Investigation, Methodology, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Córdova-Agurto JS**: Conceptualization, Data Curation, Investigation, Methodology, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Fiestas J**: Conceptualization, Data Curation, Investigation, Methodology, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Rojas-Vilar FJ**: Conceptualization, Data Curation, Investigation, Methodology, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Culquichicón C**: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The Peru Infectious Diseases Epidemiology Research Training Consortium (D43 TW007393), awarded by the Fogarty International Center of the US National Institutes of Health, sponsored Drs. Valladares and Culquichicon's work. The Universidad Continental funded the editorial expenses.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Valladares-Garrido MJ, Mejia CR, Rojas-Alvarado AB *et al.* **Factors associated with producing a scientific publication during medical training: evidence from a cross-sectional study of 40 medical schools in Latin America [version 2; peer review: 2 approved]** F1000Research 2022, **9**:1365 https://doi.org/10.12688/f1000research.26596.2

First published: 24 Nov 2020, **9**:1365 https://doi.org/10.12688/f1000research.26596.1

REVISED **Amendments from Version 1**

We have addressed English language issues and revised some terms and phrases. In particular, there are phrases that were conveyed more simply for a broader audience (e.g., clinical educators). In addition, we detailed the statistical analysis more precisely. The results and discussion section were also revised. All discussion paragraphs have a new sentence explaining the significance and implications of our findings. We clarified the limitations of the study and added some illustrative examples for the reader's convenience.

The Universidad Continental was listed as an affiliation due to an initial commitment to support MJVG and CC on editorial expenses, which wasn't complete due to external factors. Currently, Universidad Privada Norbert Wiener is the supporting institution of MJVG, and CC, which may take full credit due to its institutional effort on publication fees, and career development.

Any further responses from the reviewers can be found at the end of the article

Introduction

Producing a scientific publication during medical training is key to promoting continuing medical education and encouraging trainees to create cutting-edge knowledge. In doing so, students will develop research and critical thinking skills and will carry out evidence-based practice and patient-centered care with an enduring vision for pursuing a scientific career[1–3]. Latin American universities are progressively recognizing the critical importance of fostering science at the beginning of the bachelor's degree, and are implementing research-oriented courses such as research design methods, biostatistics, epidemiology, and a research-focused thesis[4]. However, there are still gaps in Latin America compared to university research systems in developed countries in terms of number of publications, quality of published articles, dissemination of studies, and funding opportunities[5]. Studies in Colombia and Brazil show that medical students consider scientific research as an important aspect of their training and that the low scientific output is influenced by the lack of inspiring and committed mentors as role models for the beginning of the scientific career[6,7]. Between 1997 and 2010, there was an 8.4% increase in student participation in manuscripts published in journals indexed in Scielo-Peru, of which 42% reported being affiliated with a medical student scientific society[4,8].

In Peru, the progress of undergraduate medical research has been strongly promoted by the Peruvian Scientific Society of Medical Students (SOCIMEP, by its acronym in Spanish), an organization that has been improving the research training of medical students for 27 years[9]. SOCIMEP is organized in scientific and academic committees and is made up of 38 local scientific societies in all Peruvian medical schools. This society is recognized for the organization of international, national, and local scientific conferences[9]. SOCIMEP also encourages the active participation of societies and integrates them into a nationwide research network, and provides connections to experienced research mentors. Being affiliated to a local scientific society affiliated to SOCIMEP is associated with a higher scientific production (PR: 2.41; 95% CI: 1.55-3.74)[10]. However, only 10% of the projects carried out in local scientific societies are published in indexed journals due to poor methods applied in the studies, lack of knowledge of the editorial process, few local mentors, and lack of financial support from public agencies and institutions[11]. Funding opportunities for medical students are scarce in local medical schools in Peru and in much of Latin America. Overall government investment is disproportionately granted and often contradictory to local public health needs, detracting from the importance of well-implemented laboratories and full-time research-focused faculties[12]. In Peru, less than 30% of universities have funding programs for students to conduct thesis research, or awards for student research programs[13].

The promotion of science among medical students in Latin America is a multi-sectoral issue that is hampered by governments' lack of knowledge about investing in well-structured national research and innovation systems, as well as by the lack of support from local universities, their lack of investment in research facilities, and the lack of mentors with international research experience[3]. Improvement of the scientific system in Latin America might be valuable for other regions of the world, by promoting high-quality research at the undergraduate-level, an integrative ecosystem of research and education would be consolidated for to better medical practice, and training of health professionals in similar settings across the globe. In this scenario, our study aims to determine the factors associated with scientific publication during medical training, in order to identify the needs of local Latin American scientific societies for the implementation of continuing education programs in research. Our hypothesis is that there are factors in medical training associated with the production of a scientific publication during undergraduate training.

Methods

Study design

This is a secondary data analysis of a cross-sectional study initially conducted in 2016 to assess the use of information and communication technologies (ICTs) in medical students across Latin America[14,15]. This study evaluated 40 medical student scientific societies in Latin America. The self-report of having a scientific publication was evaluated. In addition, the following variables were used to explore factors associated with scientific publication: gender, age, university, current year of study, medical student scientific society membership, English proficiency, previous career studied, courses in scientific databases including PubMed, Scopus and Scielo, courses in scientific writing skills, courses in scientific navigation, courses in Zotero, use of Sci-Hub, and access to and provision of pirated academic accounts.

Population and sampling

The primary study surveyed 11,587 students from 40 medical schools, including two from Ecuador, two from Panama, four from Paraguay, three from Bolivia, 18 from Peru, two from Mexico, two from Venezuela, one from Honduras, three from Colombia, one from Chile and two from Argentina. Medical students enrolled in the 2016-I term were included in this study and those doing their internship were excluded.

We performed stratified sampling using the academic year in medical school as the stratum. The estimated sample size

for each research center was 289 medical students, to which 10% was added to allow for dropouts. Thus, we set out to survey 318 medical students at each university. We considered a sample size calculation with 80% power and 5% significance for an infinite sample size. The use of these parameters is a convention to determine a conservative sample size that can detect a minimal difference for the outcome. As for the selection of the participants, the interview team went into the course with the highest credit in each academic year and chose the students who were seated in an odd place per row. In three universities, the sample size was not large enough to reach the minimum required, so we surveyed to all students.

Operational procedures

In 2015, the ICTs project was awarded an amount of money for publication in the multicenter project competition of the 30th International Congress of the Latin American Federation of Scientific Societies of Medical Students (FELSOCEM). This award allowed the authors to contact the researchers of the FELSOCEM international collaboration network for the development of the study. We were able to register teams from 40 out of 69 Scientific Societies of Medical Students (SOCEM) throughout Latin America. Each scientific society had at least one team with three medical students who received training on scientific integrity[16], standardized methods for survey participants, data entry procedures and quality control of the datasets.

In each medical school, a designated team of interviewers surveyed at the beginning or end of lectures, prioritizing that students had enough time for their comfort. The questionnaire was given to each selected student after explaining the objective of the study and the duration of the survey (approximately 15 minutes). The survey was self-reported, that is, the participants provided the answers themselves. An English translation of the survey is available as *Extended data*[17].

Measures

Self-reporting of manuscript publication was analyzed as a binary outcome. Multinomial variables included gender, age, current year of degree, English proficiency, courses in PubMed, courses in Scopus, courses in Scielo, and provision of pirated scholarly accounts. Binary variables included university, affiliation with a medical student scientific society, previously studied career, scientific database courses, scientific writing courses, scientific navigation courses, Zotero courses, Sci-hub usage, and pirated academic account usage. All these variables were self-reported.

Sci-hub usage is defined as the use of the web service to read and download restricted scientific articles that are typically paid or subscription-linked at academic institutions. Use of pirated academic accounts is the use of any account provided by a teacher, student or other person that helps the student find and download articles from academic institutions that subscribe to scientific journals or databases.

Data analysis

The association between self-reporting of manuscript publication and its covariates was assessed using chi-square tests for categorical variables and the Mann-Whitney U test for numerical variables. Poisson family regressions were performed using a log link function and mixed effects multilevel models. Nested models were estimated following a forward manual selection method using likelihood ratio tests. Covariates with significant p-values ($p < 0.05$) were included in the further nested model until statistical significance was not reached. This method was used to obtain a parsimonious multivariate model, which retains the least amount of covariates to explain the variance of the outcome. Crude and adjusted prevalence ratios (PR) were estimated with 95% confidence intervals (95%CI). All hypotheses were tested with a significance of 5%. The analysis was performed using Stata 15.1. The code is openly available on GitHub and Zenodo[18].

Ethical considerations

This study was classified as minimal risk for participants by the Institutional Review Board of San Bartolome's Hospital (CIE15325-15), and issued its approval. Trained interviewers obtained verbal consent from participants and provided them with an anonymous self-administered survey. Each survey was assigned a numerical ID to protect the privacy of the participants.

Results

A total of 11,587 medical students completed the survey. The mean age was 21±2.9 years, 53% were female, 12.5% (n=1,449) were affiliated with a medical student scientific society, and 14.1% (n=1,618) reported advanced English language skills. The individual-level responses are available as *Underlying data*[19].

Scientific writing courses were attended by 65.1% (n=3,989) of the students, and 7.9% (n=893) had published at least one scientific article during their medical training. Out of 6,632 students, 19.2% (n=1,273) used Sci-Hub at some point in their career (Table 1).

There were differences in the prevalence of scientific publications among first- and final-year medical students (4.3% first year vs. 13% final year), membership in a medical student scientific society (12.43% yes vs. 7, 24% no), advanced and elementary English proficiency (11.2% advanced vs. 6.4% elementary), completion of a scientific writing course (14.6% yes vs. 4.3% no), use of Sci-Hub (19.3% yes vs. 4.7% no) and possession of pirated academic accounts (15.3% yes vs. 5.5% no) (Table 2).

The nested models progressively selected the following covariates: scientific writing courses, pirated academic accounts, universities, Zotero courses, scientific database courses, year of study, previous degree, English proficiency, and medical student scientific society membership. The prevalence of having a scientific publication was 85% (pPR=1.85, 95% CI=1.59–2.15, p<0.001) higher in students who took a scientific

Table 1. Characteristics of medical students from 40 schools of medicine in Latin America.

Characteristics	N=11,587	n	%
Gender	11,587		
Male		5,363	46.3
Female		6,224	53.7
Age (years)*		21±2.86	
University	11,587		
National		6,119	52.8
Private		5,468	47.2
Current year of career	11,586		
1st		2,575	22.2
2nd		2,486	21.5
3rd		2,053	17.7
4th		1,969	17.0
5th		1,585	13.7
6th		918	7.9
Affiliated to a Scientific Medical Student Society	11,587		
No		10,138	87.5
Yes		1,449	12.5
English proficiency	11,499		
Elementary		2,028	17.6
Basic		4,666	40.6
Intermediate		3,187	27.7
Advanced		1,618	14.1
Studied previous career	11,574		
No		10,689	92.4
Yes		885	7.7
Courses in scientific databases	11,448		
No		5,300	46.3
Yes		6,148	53.7
Courses in PubMed	11,297		
Do not use the database		4,529	40.1
No		3,686	32.6
Yes		3,082	27.3
Courses in Scopus	11,139		
Do not use the database		9,334	83.8
No		896	8.0
Yes		909	8.2
Courses in Scielo	11,200		
Do not use the database		4,918	43.9
No		4,165	37.2
Yes		2,117	18.9
Courses in scientific writing	11,417		
No		7,428	65.1
Yes		3,989	34.9
Courses in scientific searches	11,458		
No		4,564	39.8
Yes		6,894	60.2
Courses in Zotero	11,408		
No		9,485	83.1
Yes		1,923	16.9
Use of Sci-Hub	6,632		
No		5,359	80.8
Yes		1,273	19.2
Pirated academic accounts	11,136		
No		8,622	77.4
Yes		2,514	22.6
Provider of pirated academic accounts	11,063		
Student		1,751	15.8
Professor		1,817	16.4
Both		14	16.4
Do not answer		7,481	67.6
Scientific publication	11,316		
No		10,423	92.1
Yes		893	7.9

* Mean ± standard deviation.

writing course, 81% (pPR=1.81, 95%CI=1.50–2.20, p<0.001) higher for students who used Sci-Hub, and 108% (pPR=2.08, 95%CI=1.83–2.36, p<0.001) higher among students who had a pirated academic account (Table 3). Information about medical schools in Latin America is available as *Extended data*[20].

Discussion

Pirated academic accounts and use of Sci-Hub

Sci-Hub use was reported by 19.2% (n=1273) of the students surveyed, of whom 19.3% (n=243) published a manuscript during their medical training. Awareness and use of Sci-Hub

Table 2. Characteristics of medical students among scientific publication from 40 schools of medicine in Latin America.

Characteristics	Scientific publication				*P* value
	No		Yes		
	n	%	n	%	
Gender					0.529
Male	4,823	91.9	423	8.1	
Female	5,600	92.3	470	7.7	
Age (years)*	0.99				<0.001
University					<0.001
National	5,389	90.2	589	9.9	
Private	5,034	94.3	304	5.7	
Current year of career					<0.001
1st	2,417	95.7	108	4.3	
2nd	2,258	92.9	172	7.1	
3rd	1,855	92.3	155	7.7	
4th	1,763	90.9	177	9.1	
5th	1,382	89.1	169	10.9	
6th	748	87.0	112	13.0	
Affiliated to a Scientific Medical Student Society					<0.001
No	9,176	92.8	716	7.24	
Yes	1,247	87.6	177	12.43	
English proficiency					<0.001
Elementary	1,869	93.6	127	6.4	
Basic	4,240	93.1	315	6.9	
Intermediate	2,869	91.4	271	8.6	
Advanced	89	88.8	175	11.2	
Studied previous career					<0.001
No	9,684	92.7	765	7.3	
Yes	729	85.2	127	14.8	
Courses in scientific databases					<0.001
No	5,024	96.3	195	3.7	
Yes	5,356	88.5	697	11.5	
Courses in PubMed					<0.001
Do not use the database	4,220	94.2	259	5.8	
No	3,283	91.2	317	8.8	
Yes	2,720	89.8	309	10.2	
Courses in Scopus					<0.001
Do not use the database	8,531	92.7	673	7.3	
No	795	89.6	92	10.4	
Yes	798	88.4	105	11.6	
Courses in Scielo					<0.001
Do not use the database	4,624	95.3	226	4.66	
No	3,634	88.5	472	11.5	
Yes	1,905	91.5	178	8.55	
Courses in scientific writing					<0.001
No	6,993	95.7	317	4.34	
Yes	3,366	85.4	574	14.57	
Courses in scientific browsing					<0.001
No	4,298	95.4	206	4.57	
Yes	6,091	89.9	684	10.1	
Courses in Zotero					<0.001
No	8,774	93.9	570	6.1	
Yes	1,579	83.2	320	16.85	
Use of Sci-Hub					<0.001
No	5,026	95.3	246	4.67	
Yes	1,016	80.7	243	19.3	
Pirated academic accounts					<0.001
No	8,055	94.5	468	5.49	
Yes	2,102	84.7	380	15.31	
Provider of pirated academic accounts					<0.001
Student	1,486	86.1	242	3.28	
Professor	1,453	80.3	239	13.86	
Both	12	85.7	357	19.72	
Do not answer	7,140	96.7	2	14.29	

* Simple logistic regression. Beta and p-value.

Table 3. Associated factors with scientific publication among medical students from 40 schools of Medicine in Latin America and the Caribbean.

Parameters	Scientific publication															Models**
	Simple regression (1)					Multiple regression parsimonious model (2)					Adjusted parsimonious model (2) *					
	PRc	95% CI			*P* value	PRp	95% CI			*P* value	PRp	95% CI			*P* value	
Gender																
Male	Ref.										Ref.					1
Female	0.96	0.85	-	1.09	0.529						0.99	0.87	-	1.12	0.865	
Age (years)*											1.02	0.99	-	1.045	0.305	2
University																
National	1.73	1.51	-	1.98	<0.001	1.82	1.58	-	2.09	<0.001						
Private	Ref.					Ref.										
Current year of career																
1st	Ref.					Ref.										
2nd	1.65	1.31	-	2.09	<0.001	1.27	1.01	-	1.60	0.043						
3rd	1.80	1.42	-	2.29	<0.001	1.39	1.09	-	1.76	0.007						
4th	2.13	1.69	-	2.69	<0.001	1.29	1.02	-	1.64	0.036						
5th	2.55	2.02	-	3.22	<0.001	1.64	1.30	-	2.07	<0.001						
6th	3.04	2.36	-	3.92	<0.001	1.79	1.38	-	2.33	<0.001						
Affiliated to a Scientific Medical Student Society																
No	Ref.					Ref.										
Yes	1.72	1.47	-	2.00	<0.001	1.36	1.16	-	1.59	<0.001						
English proficiency																
Elementary	Ref.					Ref.										
Basic	1.09	0.89	-	1.33	0.412	1.11	0.91	-	1.35	0.302						
Intermediate	1.36	1.11	-	1.66	0.003	1.22	0.99	-	1.49	0.057						
Advanced	1.76	1.42	-	2.19	<0.001	1.51	1.21	-	1.87	<0.001						
Studied previous career																
No	Ref.					Ref.										
Yes	2.03	1.70	-	2.41	<0.001	1.68	1.41	-	2.00	<0.001						
Courses in scientific databases																
No	Ref.					Ref.										
Yes	2.21	1.90	-	2.57	<0.001	1.58	1.33	-	1.88	<0.001						

Parameters	Scientific publication															Models**
	Simple regression (1)					Multiple regression parsimonious model (2)					Adjusted parsimonious model (2) *					
	PRc	95% CI			*P* value	PRp	95% CI			*P* value	PRp	95% CI			*P* value	
Courses in PubMed																3
Do not use the database	Ref.											Ref.				
No	1.52	1.30	-	1.78	<0.001						0.96	0.81	-	1.12	0.582	
Yes	1.76	1.51	-	2.07	<0.001						0.86	0.72	-	1.01	0.069	
Courses in Scopus																4
Do not use the database	Ref.										Ref.					
No	1.42	1.15	-	1.74	0.001						1.02	0.82	-	1.27	0.848	
Yes	1.59	1.31	-	1.93	<0.001						0.92	0.76	-	1.13	0.427	
Courses in Scielo																5
Do not use the database	Ref.										Ref.					
No	2.47	2.12	-	2.87	<0.001						1.47	1.24	-	1.74	<0.001	
Yes	1.83	1.52	-	2.22	<0.001						0.95	0.77	-	1.18	0.644	
Courses in scientific writing																
No	Ref.					Ref.										
Yes	3.36	2.95	-	3.83	<0.001	1.85	1.59	-	2.15	<0.001						
Courses in scientific searches																
No	Ref.															
Yes	3.08	2.64	-	3.60	<0.001											
Courses in Zotero																
No	Ref.					Ref.										
Yes	2.76	2.43	-	3.14	<0.001	1.66	1.45	-	1.90	<0.001						
Use of Sci-Hub																6
No	Ref.										Ref.					
Yes	4.14	3.50	-	4.88	<0.001					1.81	1.50	-	2.20	<0.001		
Pirated academic accounts																
No	Ref.					Ref.										
Yes	2.79	2.45	-	3.17	<0.001	2.08	1.825	-	2.36	<0.001						

Parameters	Scientific publication													Models**
	Simple regression (1)					Multiple regression parsimonious model (2)				Adjusted parsimonious model (2) *				
	PRc	95% CI			*P* value	PRp	95% CI			*P* value	PRp	95% CI			*P* value	
Provider of pirated academic accounts																
Student	4.23	3.56	-	5.01	<0.001						3.86	3.05	-	4.90	<0.001	7
Professor	6.02	5.15	-	7.02	<0.001					4.56	3.85	-	5.39	<0.001		
Both	4.36	1.20	-	15.82	0.025						3.83	0.93	-	15.69	0.062	
Do not answer	Ref.										Ref.					

(1) Poisson´s regression model with robust variance.
(2) Poisson´s regression model with robust variance and multilevel analysis.
* Multiple regression parsimonious model was indepently adjusted by each variable below
Abbreviations: PRc, Crude prevalence ratio; PRp, Parsimonious model's prevalence ratio; PRa, adjusted parsimonious' prevalence ratio.

may be due to the strong need for access to high-level scientific evidence behind a paywall. This need is often reinforced because many medical schools do not offer access to high quality scientific journals or databases. However, medical students have reported difficulties in accessing Sci-Hub because it is considered an illegal service in many regions, meaning that the web domain is often blocked[21–24].

Sci-Hub use was associated with a higher prevalence of scientific publication among medical students (PR: 1.81; 95%CI: 1.50-2.20). Students feel a strong need of access to paid articles, leading them to seek free access on Sci-Hub[23,25]. However, even those students who do not face a paywall, found using Sci-Hub reduced the time and increased simplicity of browsing[26]. In addition, many researchers and students identify Sci-Hub as a faster option that is not limited to their institution's catalog[24]. This process of rapid acquisition of scientific articles offered by sci-hub is probably homogeneous among high- and low-income countries around the world[27]. More than 56,000 article downloads through Sci-Hub come from different cities on the east coast of the United States, especially from cities where major universities subscribe to different publishers[26].

Use of pirated academic accounts was associated with a higher prevalence of scientific publishing (PR: 2.08; 95% CI: 1.83–2.36). Pirated accounts are an alternative to institutional licenses for obtaining access to journals, books or specialized databases such as Scopus or Web of Science. Although fee-based services are financed by governmental institutions in low- and middle-income countries (LMICs), they are not widely distributed or have not been applied in LMICs[28]. Other alternatives, e.g., HINARI, allow access to fee-based articles in LMICs, but is available to the academic and research community only from certified institutions that have reached certain milestones defined by local science systems[29]. This complex context leads users to exchange, lend or acquire access accounts or proxy links to institutional journal catalogs under non-legal terms[27].

Courses in scientific writing

One third (34.9%) of the students who have published a scientific publication have attended a course on scientific writing. Attending a scientific writing course increased the prevalence of scientific publications by 85% (pPR=1.85, 95%CI=1.59–2.15, p<0.001). This is likely due to the great need for medical students to improve their skills to effectively communicate scientific findings, make relevant scholarly reflection and increase the chances of being accepted into a scientific journal[30]. New medical students in research training are eager to train in scientific writing skills and are looking for an experienced mentor to train them[31]. When research courses are not locally available in institutions, students are likely to seek training in online short-courses, for example, the Brazilian initiative DivulgaMicro was a course funded by the Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP) to train early-career researchers to translate complex scientific messages into understandable pieces of information for members of the scientific community[32]. Within 30 days of its launch, the website registered 1,026 users from different regions of the world, including Latin America, the United Kingdom, Pakistan, Germany and Canada. This is one of the most visited free and open science communication workshops, in which more than 600 junior medical student researchers were trained[32]. In this context, our results may encourage medical education programmers to implement scientific writing courses as part of their curricula.

English proficiency

An advanced level of English was achieved in 14.1% of the students, and 11.2% published a scientific manuscript

during their medical training. In addition, the prevalence of scientific publications increased by 51% among students with advanced English proficiency (pPR=1.51, 95%CI=1.21–1.87, p<0.001). Students are encouraged to understand scientific evidence written in English[33]. TOEFL score correlates with publication in a medical journal (correlation coefficient: 0.63)[34]. Scientific journals preferentially accept articles from native English speakers over non-native English speakers (acceptance rate 7% vs 3.6%, respectively)[35]. Also, Americans are 49% more likely to have an article reviewed or accepted in a U.S. journal compared to non-native speakers[35]. Second- to sixth-year medical students who attended a training course on scientific writing in English reported that 53% of them perceived that they were not proficient enough in English to publish a manuscript in English-language journals[36]. Our findings may help medical educators to train student researchers to read and write more articles in English. In addition, medical curricula could establish effective medical English courses to help improve the rate of scientific publications.

The association between scientific publications and advanced English proficiency could be due to students' desire to pursue an academic education abroad offered by institutions that demand academic excellence and high potential. In 2016, the Peruvian Program for Scholarships and Educational Credits (PRONABEC) jointly funded Fulbright, FONDECYT, and Chevening scholarships in Peru that benefited 14, 6 and 15 Peruvian graduate applicants, respectively[37]. In this way, scholarship recipients could be trained at leading foreign universities, producing a generation of researchers with master's and doctoral degrees who, upon returning to their countries of origin, seek to improve the science and technology system[38–40]. During 2004-2012, the Fogarty International Clinical Research Scholars and Fellows Program funded promising initiatives by students from low- and middle-income countries with English proficiency, whose scientific discoveries may address long-term global health needs[41,42]. This approach has become Fogarty's hallmark: bringing great science to solve local problems with global reach and building local research capabilities[42]. During 2014-2015, Fogarty has contributed substantially to the training of more than 6,100 global health leaders, 140 of whom have earned doctoral degrees in epidemiology and 96 in public health[43]. The process of obtaining scholarships and future academic degrees could be improved by equalizing opportunities at the undergraduate level.

Fogarty International Center bridges U.S. National Institutes of Health with global health research community; 85-90% of trained fellows return to LMICs and obtain research positions in universities, government agencies, and institutes[42]. However, young Latin American researchers and foreign-trained postdoctoral researchers face difficulties due to an unfavorable scientific system[29]. For example, the Peruvian administration's investment in the advancement of science and research is still insufficient, at only 0.12% of gross domestic product compared to 0.36% in Chile, 1.3% in Brazil, and 2.8% in the United States[44,45]. This is a concerning situation that must be addressed at the political level to efficiently solve public health needs.

Medical student scientific society membership

Our results showed that membership in a medical student scientific society increased the prevalence of scientific publication by 36% (pPR:1.36, 95%CI=1.16-1.59, p<0.01). Student scientific societies, such as SOCIMEP, attempt to fill the gaps in research training and provide students with the mentors, courses and scientific opportunities to pursue a research career[9,46]. With more than 30 years of operations with local scientific societies throughout Peru, SOCIMEP promotes regional, national and local research events (CUMIS), annual scientific congresses and foundation courses in epidemiology, research design, and biostatistics[47]. SOCIMEP's overall reach was reflected in the 242 articles published by scientific societies, of which 11% (n=67) were published in Q1 journals, under the tutelage of highly experienced national researchers[48]. SOCIMEP's presence in Peru demonstrates the importance of an integrated institution that could not only equalize opportunities for students, but also improve scientific production in the country. Our results suggest that this student research system could be an effective model for other similar contexts.

Limitations

Our results have limitations that are described in the following statements. First, several questionnaire items were self-reported, which may cause outcome misclassification. This means that a participant is classified to the wrong group, e.g., a student who is proficient in English feels unskilled and their response leads them being classified as a non-proficient student. and increase the potential of information bias. However, we tried to control this situation by motivating the students to answer the questionnaire in an honest manner and not to rush them; in this sense, our result is consistent with reality. Second, all 40 medical schools were affiliated with FELSOCEM, which indicates a possible selection bias because this Latin American institution is integrated by medical schools that meet standardized parameters of undergraduate scholarly. Therefore, our results are useful for these schools but should be extended to other similar local and regional realities in different countries. Third, some other factors may be missing to better understand the medical training characteristics that may influence scientific publication. For example, the type of university (private or public), the gross national income devoted to research in each participating country, and the presence of highly qualified researchers in medical schools. However, this study provides relevant information to design new studies addressing the scientific production of medical students.

Conclusion

Factors associated with producing a scientific publication in medical students during their medical training in Latin America are being affiliated to a scientific society of medical students, having an advanced command of English, having

attended a scientific writing course, the use of Sci-Hub and the use of pirate accounts. The promotion of science among medical students in Latin America is a multisectoral issue. Its development must be addressed as part of multilevel strategies coming from the highest governmental authorities. In this way, universities would be empowered and a committed scientific system would be built in each nation.

Data availability

Underlying data

Figshare: Scientific article. https://doi.org/10.6084/m9.figshare.13061699.v2[19].

This project contains the underlying data in DTA and CSV formats.

Extended data

Figshare: Technological and educational factors associated with the use of information sources in medical students from Latin America. https://doi.org/10.6084/m9.figshare.13070603.v1[17].

This project contains an English-language copy of the questionnaire used for data collection.

Figshare: Latin American medical students surveyed in 2016 - Supplementary materials from a cross-sectional study of 40 medical schools surveyed in Latin America. https://doi.org/10.6084/m9.figshare.13070693.v1[20].

This project contains a list of the medical schools surveyed for this study.

Analysis code used in this study is available at: https://github.com/culquichicon/Scientific_writing.

Archived code at time of publication: https://doi.org/10.5281/zenodo.3730359[18].

Analysis code license: GNU General Public License v3.0.

References

1. Riggs KR, Reitman ZJ, Mielenz TJ, *et al.*: **Relationship Between Time of First Publication and Subsequent Publication Success Among Non-PhD Physician-Scientists.** *J Grad Med Educ.* 2012; **4**(2): 196–201.
PubMed Abstract | Publisher Full Text | Free Full Text
2. Gouda MA, Zidan HS, Marey AA, *et al.*: **Medical undergraduates' contributions to publication output of world's top universities in 2013.** *QJM.* 2016; **109**(9): 605–11.
PubMed Abstract | Publisher Full Text
3. Corrales-Reyes IE, Dorta-Contreras AJ: **Students scientific production: a proposal to encourage it.** *Medwave.* 2018; **18**(1): e7166.
PubMed Abstract | Publisher Full Text
4. Taype-Rondán Á, Lajo-Aurazo Y, Gutiérrez-Brown R, *et al.*: **Aporte de las sociedades estudiantiles en la publicación científica en Scielo-Perú, 2009 - 2010.** *Rev Peru Med Exp Salud Publica.* 2011; **28**(4): 691–2.
Publisher Full Text
5. Machado-Alba J, Machado-Duque M: **The Role of Research Incubators in Encouraging Research and Publication Among Medical Students.** *Acad Med.* 2014; **89**(7): 961–2.
PubMed Abstract | Publisher Full Text
6. Bonilla-Escobar FJ, Bonilla-Velez J, Tobón-García D, *et al.*: **Medical student researchers in Colombia and associated factors with publication: a cross-sectional study.** *BMC Med Educ.* 2017; **17**(1): 254.
PubMed Abstract | Publisher Full Text | Free Full Text
7. Moraes DW, Jotz M, Menegazzo WR, *et al.*: **Interest in research among medical students: Challenges for the undergraduate education.** *Rev Assoc Med Bras (1992).* 2016; **62**(7): 652–8.
PubMed Abstract | Publisher Full Text
8. Huamaní C, Chávez-Solis P, Mayta-Tristán P: **Aporte estudiantil en la publicación de artículos científicos en revistas médicas indizadas en Scielo-Perú, 1997 - 2005.** *An Fac Med.* 2008; **69**(1): 42–5.
Publisher Full Text
9. Cvetkovic-Vega A, Inga-Berrospi F, Abel Mestas C: **Organizaciones científicas estudiantiles como semilleros de líderes y gestores de la investigación científica en el Perú: SOCIMEP.** *Acta Méd Peru.* 2017; **34**(1): 70–71.
Publisher Full Text
10. Toro-Huamanchumo CJ, Failoc-Rojas VE, Díaz-Vélez C: **Participación en sociedades científicas estudiantiles y en cursos extracurriculares de investigación, asociados a la producción científica de estudiantes de medicina humana: estudio preliminar.** *FEM Rev Fund Educ Médica.* 2015; **18**(4): 293–8.
Publisher Full Text
11. Toro-Polo M, Pereyra-Elías R, Nizama-Vía A, *et al.*: **Publicación de los trabajos presentados a los congresos científicos de estudiantes de medicina, Perú 2002-2009: características y factores asociados.** *Rev Peru Med Exp Salud Publica.* 2012; **29**(4): 461–8.
PubMed Abstract | Publisher Full Text
12. Moses H 3rd, Matheson DHM, Cairns-Smith S, *et al.*: **The anatomy of medical research: US and international comparisons.** *JAMA.* 2015; **313**(2): 174–89.
PubMed Abstract | Publisher Full Text
13. Toro-Huamanchumo CJ, Arce-Villalobos LR, Gonzales-Martínez J, *et al.*: **Financiamiento de la investigación en pregrado en las facultades de medicina peruanas.** *Gac Sanit.* 2017; **31**(6): 541–2.
PubMed Abstract | Publisher Full Text
14. Mejia CR, Valladares-Garrido MJ, Quintana-Gomez S, *et al.*: **Carrera previa como factor asociado al uso de buscadores científicos entre estudiantes de medicina latinoamericanos: cuando la experiencia no cuenta.** *Educ Médica.* 2019; **20**(Supplement 1): 131–5.
Publisher Full Text
15. Mejia CR, Valladares-Garrido MJ, Almanza-Mio C, *et al.*: **Participación en una sociedad científica de estudiantes de Medicina asociada a la producción científica extracurricular en Latinoamérica.** *Educ Médica.* 2019; **20**(Supplement 1): 99–103.
Publisher Full Text
16. QUIPU- Centro Andino de Investigación y Entrenamiento en Informática para la Salud Global de la Universidad Peruana Cayetano Heredia: **Conducta Responsable en Investigación**. 2018; [accessed 17 january 2020].
Reference Source
17. Culquichicón C: **Technological and educational factors associated with the use of information sources in medical students from Latin America.** *figshare.* Journal contribution. 2020.
http://www.doi.org/10.6084/m9.figshare.13070603.v1
18. Culquichicón C: **culquichicon/Scientific_writing: Associated factors with reaching a scientific publication during medical training: evidence from 40 medical schools surveyed in Latin America. (Version sw_v1)**. *Zenodo.* 2020.
http://www.doi.org/10.5281/zenodo.3730359
19. Culquichicón C, Valladares-Garrido MJ: **Scientific article.** *figshare.* Dataset. 2020.
http://www.doi.org/10.6084/m9.figshare.13061699.v2
20. Culquichicón C: **Latin American medical students surveyed in 2016 - Supplementary materials from a cross-sectional study of 40 medical schools surveyed in Latin America.** *figshare.* Journal contribution. 2020.
http://www.doi.org/10.6084/m9.figshare.13070693.v1
21. Mejia CR, Valladares-Garrido MJ, Miñan-Tapia A, *et al.*: **Use, knowledge, and perception of the scientific contribution of Sci-Hub in medical students: Study in six countries in Latin America.** *PLoS One.* 2017; **12**(10): e0185673.
PubMed Abstract | Publisher Full Text | Free Full Text

22. Schiermeier Q: **Pirate research-paper sites play hide-and-seek with publishers.** *Nature.* 2015.
Publisher Full Text

23. Himmelstein DS, Romero AR, Levernier JG, *et al.*: **Sci-Hub provides access to nearly all scholarly literature.** *eLife.* 2018; **7**: e32822.
PubMed Abstract | Publisher Full Text | Free Full Text

24. Hoy MB: **Sci-Hub: What Librarians Should Know and Do about Article Piracy.** *Med Ref Serv Q.* 2017; **36**(1): 73–8.
PubMed Abstract | Publisher Full Text

25. Greshake B: **Looking into Pandora's Box: The Content of *Sci-Hub* and its Usage [version 1; peer review: 2 approved, 2 approved with reservations].** *F1000Res.* 2017; **6**: 541.
PubMed Abstract | Publisher Full Text | Free Full Text

26. Bohannon J: **Who's downloading pirated papers? Everyone.** *Science.* 2016; **352**(6285): 508–12.
PubMed Abstract | Publisher Full Text

27. Till BM, Rudolfson N, Saluja S, *et al.*: **Who is pirating medical literature? A bibliometric review of 28 million Sci-Hub downloads.** *Lancet Glob Health.* 2019; **7**(1): e30–1.
PubMed Abstract | Publisher Full Text

28. Danda D: **Cost of publication - who pays for it?** *Int J Rheum Dis.* 2014; **17**(4): 358.
PubMed Abstract | Publisher Full Text

29. Ciocca DR, Delgado G: **The reality of scientific research in Latin America; an insider's perspective.** *Cell Stress Chaperones.* 2017; **22**(6): 847–52.
PubMed Abstract | Publisher Full Text | Free Full Text

30. Bhardwaj P, Sinha S, Yadav RK: **Medical and scientific writing: Time to go lean and mean.** *Perspect Clin Res.* 2017; **8**(3): 113–117.
PubMed Abstract | Free Full Text

31. Kennedy AB: **Journal Aspirations: Improving Scientific Writing and Publication Through a Writing Mentorship Program.** *Int J Ther Massage Bodyw.* 2017; **10**(2): 1–2.
PubMed Abstract | Publisher Full Text | Free Full Text

32. Oliveira LMA, Bonatelli ML, Pinto TCA: ***DivulgaMicro*: A Brazilian Initiative To Empower Early-Career Scientists with Science Communication Skills.** *J Microbiol Biol Educ.* 2019; **20**(1): 7.
PubMed Abstract | Publisher Full Text | Free Full Text

33. Rezaeian M: **The pitfalls of writing and publishing a scientific health manuscript in English for non-English language scholars.** *Epidemiol Health.* 2015; **37**: e2015021.
Reference Source

34. Man JP, Weinkauf JG, Tsang M, *et al.*: **Why do some countries publish more than others? An international comparison of research funding, English proficiency and publication output in highly ranked general medical journals.** *Eur J Epidemiol.* 2004; **19**(8): 811–7.
PubMed Abstract | Publisher Full Text

35. Link AM: **US and Non-US Submissions: An Analysis of Reviewer Bias.** *JAMA.* 1998; **280**(3): 246–7.
PubMed Abstract | Publisher Full Text

36. Molina-Ordóñez J, Huamaní C, Mayta-Tristán P: **Apreciación estudiantil sobre la capacitación universitaria en investigación: Estudio preliminar.** *Rev Peru Med Exp Salud Publica.* 2008; **25**(3): 325–329.
Reference Source

37. PRONABEC: **Memoria Anual 2016.** 2018. [accessed 10 january 2020].
Reference Source

38. Fulbright Perú: **Beca de Posgrado Fulbright.** 2019. [accessed 10 january 2020].
Reference Source

39. Fondecyt: **Becas y Co-financiamiento de Concytec.** 2020. [accessed 10 january 2020].
Reference Source

40. Chevening: **English language.** 2019. [accessed 10 january 2020].
Reference Source

41. Heimburger DC, Carothers CL, Blevins M, *et al.*: **Impact of Global Health Research Training on Career Trajectories: The Fogarty International Clinical Research Scholars and Fellows Program.** *Am J Trop Med Hyg.* 2015; **93**(3): 655–61.
PubMed Abstract | Publisher Full Text | Free Full Text

42. Bridbord K, Weymouth KH, Puderbaugh A, *et al.*: **Fifty Years of Supporting Global Health Research at the NIH Fogarty International Center.** *Ann Glob Health.* 2019; **85**(1): 43.
PubMed Abstract | Publisher Full Text | Free Full Text

43. National Institutes of Health: **Report of the Director National Institutes of Health Fiscal Years 2014 and 2015.** 2018.
Reference Source

44. Cáceres CF, Mendoza W: **Globalized Research and "National Science": The Case of Peru.** *Am J Public Health.* 2009; **99**(10): 1792–8.
PubMed Abstract | Publisher Full Text | Free Full Text

45. Bajak A: **What should Peru do to improve its science?** *Nature.* 2019; **576**(7787): S65–S67.
PubMed Abstract | Publisher Full Text

46. Cvetkovic-Vega A: **Propuesta de una sociedad científica de estudiantes de medicina afiliada a SOCIMEP e IFMSA-Perú: SOCEMURP y su modelo de estructura mixta.** *FEM.* 2017; **20**(2): 89–89.
Reference Source

47. Chalco-Huamán J, Zavala-Portugal J, Andonaire-Munaico C: **Responsabilidad social en estudiantes de medicina: experiencia de una sociedad científica estudiantil peruana.** *An Fac Med.* 2016; **77**: 69–70.
Publisher Full Text

48. Taype-Rondán A, Bazán-Ruiz S, Valladares-Garrido D: **Producción científica de las sociedades científicas de estudiantes de medicina del Perú, 2002-2012.** *CIMEL.* 2013; **18**(1): 23–29.
Reference Source

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 06 October 2022

https://doi.org/10.5256/f1000research.133710.r138675

✔ **Megan Anakin**
Otago Medical School, University of Otago, Dunedin, New Zealand

Thank you for this revised version of your article. Overall, the article has been strengthened by addressing the feedback from reviewers. Thank you for addressing all of the comments and suggestion made about the first version of your article. The revisions provide the reader with enhanced descriptions of key terminology and decisions made about the study design and procedures to enable better interpretation of the results. The revisions also provide the reader with a greater understanding of the context of the study and how the problem and findings may be relevant and applicable to their context.

Competing Interests: For transparency, I am a member of the MedEdPublish Advisory Board. This review represents my view of the article and not of the Board.

Reviewer Expertise: My research area encompasses teaching, learning, curriculum, and faculty development in health professions education.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 14 April 2022

https://doi.org/10.5256/f1000research.29360.r128598

Megan Anakin
Otago Medical School, University of Otago, Dunedin, New Zealand

Review of
Factors associated with producing a scientific publication during medical training: evidence from a cross-sectional study of 40 medical schools surveyed in Latin America. DOI: 10.12688/f1000research.26596.1

General comments

The authors provide a report that addresses the aim of a secondary data analysis to better understand the factors from medical training associated with producing a scientific publication during undergraduate study. One major recommendation is to enhance the focus of the article to emphasise the relevance, importance, and implications of study and its findings to education more generally, and to students, teachers, and medical programme curriculum and resourcing, specifically.

Note: I appreciate that English might not be the authors' first language. I have made suggestions to improve the grammar of the manuscript. For example, in first sentence of the abstract, the verb 'promote' should be 'promoting'. I have not identified all instances of where subject-verb agreement needs checking. I recommend that a copy of the manuscript is shared with a colleague who can edit the manuscript for grammar or the authors access a professional editing service. Other grammar suggestion examples are:

- Introduction section, first paragraph, first sentence: The verb 'promote' makes more sense as 'promoting' and the verb 'endure' makes more sense as 'enduring'
- Introduction section, first paragraph, second sentence: The verb 'foster' makes more sense as 'fostering'
- Introduction section, first paragraph, third sentence: The verb 'comparing' makes more sense as 'compared'
- Introduction section, second paragraph, second sentence: Subject-verb agreement. Since the SOCIMEP is singular, then the verbs should be 'fosters' and 'provides'.

Specific feedback with constructive comments

Abstract

The authors succinctly outline the topic, a problem that is addressed by the methods, and present the main results. Missing from the abstract is a few sentences outlining the discussion section. Please consider outlining the key discussion points related to the significance and implications of the results for clinical educators who will be reading this article.

Introduction

Second paragraph, awkward word choice: I'm not certain what 'stands' means in this sentence. Please consider revising this long sentence into two shorter ones. The second sentence could begin: "SOCIMEP holds international,..."

The introduction provides a well-argued warrant for the study by established the local need for the study. As a reader from outside Latin America, I am wondering how this situation might be similar and different to other regions. Please consider relating this problem to the locations of readers beyond Latin America.

Methods

The aim is clearly stated at the beginning of this section and the methods presented are appropriate to address it.

Population and sampling section. Please explain the educational outcome-related evidence that was used to determine the power calculation, or if not, state the reason for why you considered a sample size calculation with an 80% of power and 5% of significance. Please describe the census-type sampling procedure or support the terminology with a reference that describes it to the reader.

Operational procedures section. First paragraph, first sentence: Please consider replacing XXX with number or word because it looks like information has been omitted to the reader, or please further explain this reference. Please specify what the ICT project was awarded or to whom the project was awarded. Please consider revising the second to last sentence of the second paragraph to explain how participants completed the surveys to provide responses selected or written by themselves and participants took about 15 minutes to complete it.

Measures section. Please clarify the term, 'self-administered survey' by providing more information or revising the sentence so it better matches the description of the survey in the operational procedures section.

Data analysis section. The analyses performed by the authors looks appropriate, however, the explanations are dense and may not be easily understood by readers of this journal. Please revise this section so that it is suitable for clinical educators who are not necessarily biomedical scientists or statisticians. Please clarify in the text below on what statistical basis the covariates were selected so the reader does not need to make inferences about the procedures. Please explain if the forward addition assessed using a LRT and if so, explain if a chi-squared distribution or something else was used. Please also describe the criterion used to include or exclude the covariate. Otherwise, the reader has to make the assumption that this information is encompassed in the 'all hypotheses…' statement. Please consider stating the following information more thoroughly and with descriptions meaningful to a clinical educator reader: "We estimated nested models following a manual forward selection method to identify covariates associated with self-reported manuscript publication until reaching a parsimonious multivariable model. These covariates were selected using likelihood ratio tests. Crude and adjusted prevalence ratios (PR) were estimated with 95% confidence intervals (CI 95%). All hypotheses were contrasted using 5% significance"

Results

First paragraph, first sentence: Please resolve the following contradiction: The methods state that participants completed surveys, however, the first sentence of the results states that interviews

were conducted. Both cannot be true.

Please consider revising the first, second, third, and fourth paragraphs of the results section so the information presented in them is not repeated in the tables of results as well. Instead, please draw the reader's attention to important or relevant proportions, relationships, differences, and similarities among the results. In the third paragraph, the percentage stated for the prevalence of scientific publications among first-year and last-year medical students (13% vs 4.3%, respectively) is backwards from the Table 2. The remaining differences need to have the two categories identified. For example, the reader does not know by reading the sentence what being affiliated to a Scientific Medical Student Society (12.43% vs 7.24%) means unless you state (12.43% yes vs 7.24% no).

Discussion

Pirated academic accounts and use of Sci-Hub section. To appreciate the similarities and differences between Sci-hub and pirated accounts, please define and explain these two factors in the methods section for the reader. To help the reader understand the significance of the results presented in the first and second paragraphs, please explain how the use Sci-Hub might contrast with access to relevant literature provided by the participating medical schools. In the second paragraph, the fifth sentence begins with 'This'. Please specify the subject at the beginning of this sentence so the reader can appreciate what might be homogeneous and better understand the point made in the final sentence in the paragraph. In the last paragraph of this section, please make links between the finding about pirated academic accounts in the first sentence and the statements that follow it. Please help the reader to understand how the statements help the reader to understand the significance and implications of this finding.

The discussion in the sections about c*ourses in writing, English proficiency, and affiliated to a scientific medical student society* do a good job of relating the findings to the local context and the literature. Please consider how the discussion can be broadened to address how your findings might be used by others to generate insights into their own local context or make suggestions about how the findings might offer insights to educators and other educational researchers about the factors from medical training associated with producing a scientific publication during undergraduate study. Since this is a medical and health professions education journal, our readers are interested in the implications for students, teachers, and medical programme curriculum and resourcing.

Limitations section. Please give an example to explain to the reader what an undifferentiated classification of the outcome and an increase in residual confusion. Please remember the readers are clinical educators who are not necessarily biomedical scientists or statisticians. Please explain the possible influence of FELSOCEM on the results and how it may bias the results. What factors might be missing in these results that might give readers further insights into the problem of producing scientific publications during medical training. Please outline a few future directions that you or other researchers might take with the findings or the study design to extend our understanding of this problem and topic

Conclusion

Please revise the very long final sentence into at least three shorter ones to help the reader

appreciate the important concluding points.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: For transparency, I am a member of the MedEdPublish Advisory Board. This review represents my view of the article and not of the Board.

Reviewer Expertise: My research area encompasses teaching, learning, curriculum, and faculty development in health professions education.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 03 May 2022

Annel Rojas Alvarado, Universidad Privada Antenor Orrego, Piura, Peru

"**Abstract**

The authors succinctly outline the topic, a problem that is addressed by the methods, and present the main results. Missing from the abstract is a few sentences outlining the discussion section. Please consider outlining the key discussion points related to the significance and implications of the results for clinical educators who will be reading this article."

Response: Thank you. A sentence on the significance and implications of the results has been added to the end of the conclusion section.

"**Introduction**

Second paragraph, awkward word choice: I'm not certain what 'stands' means in this sentence. Please consider revising this long sentence into two shorter ones. The second sentence could begin: "SOCIMEP holds international,..."""

Response: Thank you. The overall sentence was revised.

"The introduction provides a well-argued warrant for the study by established the local need for the study. As a reader from outside Latin America, I am wondering how this situation might be similar and different to other regions. Please consider relating this problem to the locations of readers beyond Latin America."

Response: Thank you. A sentence about the problem related to other contexts was added in the last paragraph (second sentence).

"**Methods**

The aim is clearly stated at the beginning of this section and the methods presented are appropriate to address it."

Response: Thank you

*"**Population and sampling section.*** Please explain the educational outcome-related evidence that was used to determine the power calculation, or if not, state the reason for why you considered a sample size calculation with an 80% of power and 5% of significance. Please describe the censustype sampling procedure or support the terminology with a reference that describes it to the reader."

Response: Thank you. The reason for the sample size calculation with the referred parameters were explained. Census-type sampling was revised to a simpler phrase.

*"**Operational procedures section.*** First paragraph, first sentence: Please consider replacing XXX with number or word because it looks like information has been omitted to the reader, or please further explain this reference. Please specify what the ICT project was awarded or to whom the project was awarded. Please consider revising the second to last sentence of the second paragraph to explain how participants completed the surveys to provide responses selected or written by themselves and participants took about 15 minutes to complete it."
Response: Thank you. "XXX" was revised to 30th. The ICTs project was awarded an amount of money for publication (details in first & second sentence). The second to last sentence of the second paragraph was revised (please refer to this part).

*"**Measures section.*** Please clarify the term, 'self-administered survey' by providing more information or revising the sentence so it better matches the description of the survey in the operational procedures section."

Response: Thank you. The term "self-administered survey" was revised

"***Data analysis section.*** The analyses performed by the authors looks appropriate, however, the explanations are dense and may not be easily understood by readers of this journal. Please revise this section so that it is suitable for clinical educators who are not necessarily biomedical scientists or statisticians. Please clarify in the text below on what statistical basis the covariates were selected so the reader does not need to make inferences about the procedures. Please explain if the forward addition assessed using a LRT and if so, explain if a chi-squared distribution or something else was used. Please also describe the criterion used to include or exclude the covariate. Otherwise, the reader has to make the assumption that this information is encompassed in the 'all hypotheses…' statement. Please consider stating the following information more thoroughly and with descriptions meaningful to a clinical educator reader: "We estimated nested models following a manual forward selection method to identify covariates associated with self-reported manuscript publication until reaching a parsimonious multivariable model. These covariates were selected using likelihood ratio tests. Crude and adjusted prevalence ratios (PR) were estimated with 95% confidence intervals (CI 95%). All hypotheses were contrasted using 5% significance""

Response: Thank you. We clarified the statistical basis for the forward selection method and other procedure details so that the information is meaningful to clinical educator readers.

"**Results**

First paragraph, first sentence: Please resolve the following contradiction: The methods state that participants completed surveys, however, the first sentence of the results states that interviews were conducted. Both cannot be true."

Response: Thank you. The term "were interviewed" was clarified to "completed the survey" to avoid the contradiction.

"Please consider revising the first, second, third, and fourth paragraphs of the results section so the information presented in them is not repeated in the tables of results as well. Instead, please draw the reader's attention to important or relevant proportions, relationships, differences, and similarities among the results. In the third paragraph, the percentage stated for the prevalence of scientific publications among first-year and last-year medical students (13% vs 4.3%, respectively) is backwards from the Table 2. The remaining differences need to have the two categories identified. For example, the reader does not know by reading the sentence what being affiliated to a Scientific Medical Student Society (12.43% vs 7.24%) means unless you state (12.43% yes vs 7.24% no)."

Response: Thank you. Relevant results were highlighted in the text. In addition, the results in Table 2 were revised and categories were included with their corresponding percentages.

Discussion

"***Pirated academic accounts and use of Sci-Hub section.*** To appreciate the similarities and differences between Sci-hub and pirated accounts, please define and explain these two factors in the methods section for the reader. To help the reader understand the significance of the results presented in the first and second paragraphs, please explain how

the use Sci-Hub might contrast with access to relevant literature provided by the participating medical schools. In the second paragraph, the fifth sentence begins with 'This'. Please specify the subject at the beginning of this sentence so the reader can appreciate what might be homogeneous and better understand the point made in the final sentence in the paragraph. In the last paragraph of this section, please make links between the finding about pirated academic accounts in the first sentence and the statements that follow it. Please help the reader to understand how the statements help the reader to understand the significance and implications of this finding."

Response: Thank you. Use of Sci-Hub and use of pirated accounts were defined and explained in the methods section. The contrast between Sci-Hub usage and access provided by medical schools is detailed in paragraph 1 of the referred section. The subject was specified in the fifth sentence of the second paragraph. Sentences in the last paragraph have been revised to better link them to the first sentence of this part.

"The discussion in the sections about courses in writing, English proficiency, and affiliated to a scientific medical student society do a good job of relating the findings to the local context and the literature. Please consider how the discussion can be broadened to address how your findings might be used by others to generate insights into their own local context or make suggestions about how the findings might offer insights to educators and other educational researchers about the factors from medical training associated with producing a scientific publication during undergraduate study. Since this is a medical and health professions education journal, our readers are interested in the implications for students, teachers, and medical programme curriculum and resourcing."

Response: Thank you. Suggestions were added at the end of each paragraph in the referred sections.

"***Limitations section.*** Please give an example to explain to the reader what an undifferentiated classification of the outcome and an increase in residual confusion. Please remember the readers are clinical educators who are not necessarily biomedical scientists or statisticians. Please explain the possible influence of FELSOCEM on the results and how it may bias the results. What factors might be missing in these results that might give readers further insights into the problem of producing scientific publications during medical training. Please outline a few future directions that you or other researchers might take with the findings or the study design to extend our understanding of this problem and topic"

Response: Thank you. Undifferentiated classification of the outcome and residual confounding were detailed. The influence of FELSOCEM was also explained. A brief list of relevant factors were added. Future directions were outlined.

"**Conclusion**

Please revise the very long final sentence into at least three shorter ones to help the reader appreciate the important concluding points."

Response: Thank you. The final sentence was revised to three shorter ones.

Competing Interests: No competing interest are present for this response

Reviewer Report 14 December 2020

https://doi.org/10.5256/f1000research.29360.r75348

Virgilio Efrain Failoc Rojas
Research Unit for the Generation and Synthesis of Evidence in Health, Saint Ignatius of Loyola University, Lima, Peru

This study brings a broad view of the current efforts of local Scientific Societies of Medical Students on promoting research among medical students. The authors found relevant factors associated with producing a scientific manuscript during medical school including the use of pirated academic accounts, training in basic research skills, English proficiency, and ultimately being affiliated to a scientific medical student society.
All of these conditions are validated by the faculties of medical schools across Latin America, and are linked between them. The outreach of this study is to give us a sight of factors to continue promoting research capacities among medical students and take advantage of the Medical Students Scientific Society as an ally in this mission.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Infectious diseases, Health interventions, Systematic Reviews and meta-analysis, Medical education.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.