Rate of Plagiarism in Personal Comments Essays of Applicants to a US Allopathic Medical School: an Observational Study

Linda S. Nield, MD[1], Manuel C. Vallejo, MD, DMD[1], Anna M. Lama, MA[1], Kristopher K. Yoon, MD[2], Alyson S. Carpenter, MA[1], Christina M. Baker, BSW[1], and Gerald R. Hobbs, PhD[1]

[1]Department of Medical Education, West Virginia University School of Medicine, Morgantown WV, USA; [2]Department of Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles CA, USA.

INTRODUCTION

Plagiarism is one type of a professionalism breach, and it has been reported to occur in the personal statements of residency applicants.[1,2] As there is no information in the literature about plagiarism in the personal comments essays of medical school applicants, the purpose of this study is to determine the rate of plagiarism in the personal comments essays of applications to an allopathic medical school in the United States (US).

METHODS

This study was approved by the West Virginia University (WVU) Institutional Review Board. Personal comments essays had to be submitted to the American Medical College Application Service (AMCAS) for the WVU School of Medicine (WVU SoM) from 2012 to 2017 for inclusion. Essays of students enrolled at WVU SoM were excluded. The plagiarism detection software, *Turnitin* (Oakland, CA, USA), which contains a database of approximately 70 billion resources and effectively detects exactly matching written content [3] was used to screen for plagiarism in the de-identified essays.

The *Turnitin* software results of 600 screened personal comments essays were reviewed by four investigators (LSN, MCV, AML, KKY) and judged as "plagiarized low" if a match of content to an existing work was less than 10%, and judged as "plagiarized high" if a match of content was 10% or greater.

STATISTICAL METHODS

The sample size determination (n1 = n2 = 300) was based on the hypothesis that the rate of plagiarism (specifically, "plagiarized high") in the first group (n1 = US applicants) would be 7.5%, and 15% in the second group (n2 = non-US applicants), using the results of residency studies as a guide.[1,2]

With a significance level of .05, two samples of 300 yield a power of approximately 83%. Pearson chi-square tests were used to compare proportions between the two groups and for continuous variables, Student's *t* test was used to compare means between the two groups, and Mann-Whitney rank sum test for ordinal data between groups. An alpha of 0.05 was used to determine statistical significance in all inferences. All statistical calculations were carried out using JMP/V14 (SAS Institute, Cary, NC).

RESULTS

The "plagiarized high" rate was 2% for both groups. There were no significant differences in the "plagiarized low" rate (22% and 25% for US and non-US respectively) and the percentages of plagiarized content per essay (Table 1). Essay contents detected as plagiarism in the "plagiarized low" category were names of organizations, titles of individuals, common sayings, and the use of quotes. Content in "plagiarized high" essays included descriptions of memorable experiences. Sources of plagiarized work originated from websites such as "samplepersonalstatement.org," "usingenglish.com," or "essayedge.com."

DISCUSSION

Our study is the first work to report the rate of plagiarism in the personal comments essays of medical school applicants. The 2% rate of plagiarism ("plagiarized high") in both US and non-US AMCAS applicants is less than the reported rates in personal statements of applicants to certain residency programs.[1,2] Parks and colleagues discovered that 13.6% of personal statements from non-US medical school graduates and 4.0% from US medical school graduates contained plagiarized material in applications to their anesthesiology residency program.[2] As a result of their study, they now screen the personal statements of every applicant and reject those found to have plagiarism. Segal et al. discovered plagiarism in 5.2% (95% CI, 4.6% to 5.9%) of statements submitted to their five

Table 1 Characteristics of Applicants and Rates of Plagiarism

	US ($n = 300$)	non-US ($n = 300$)	*P* value
Gender			0.62
Male: *n* (%)	165 (55%)	159 (53%)	–
Female: *n* (%)	135 (45%)	141 (47%)	–
Age (years ± SD)	26.5 ± 3.4	26.5 ± 3.7	0.7
Matriculated to medical school: *n* (%)	113 (38%)	55 (18%)	< 0.001
STEM major: *n* (%)	274 (91%)	291 (97%)	0.003
Highest degree obtained			0.09
None: *n* (%)	1 (0%)	1 (0%)	–
Bachelor: *n* (%)	257 (86%)	238 (79%)	–
Master: *n* (%)	41 (14%)	55 (18%)	–
PhD: *n* (%)	1 (0%)	6 (2%)	–
No. of essays with plagiarism: *n* (%)	73 (24%)	81 (27%)	0.45
Plagiarized low: *n* (%)	65 (22%)	76 (25%)	–
Plagiarized high: *n* (%)	8 (2%)	5 (2%)	–
Percent matched content: median (range)	2 (1–98)	2 (1–29)	0.24
Plagiarized low: median (range)	2 (1–7)	2 (1–7)	–
Plagiarized high: median (range)	24 (10–98)	18 (10–29)	–

STEM, science, technology, engineering, and mathematics
Plagiarized low = a match of content to an existing work was less than 10%
Plagiarized high = a match of content to an existing work was 10% or greater
Percent matched content = percent of matched content discovered per essay

largest residency programs.[2] The plagiarism rate was substantially higher for international applicants versus US applicants (14% vs. 2%, respectively).[1]

The main take away points from our study is that the rate of plagiarism detected in the personal essays of medical school applicants is relatively low when the *Turnitin* software is employed, and less in comparison with reported plagiarism in residency applications. Another unique finding is that in contrast to data reported from residency applications, there was no increased amount of plagiarism in non-US applications versus US applications. Since plagiarism has been found to be more prevalent in residency applications, our study reinforces the findings in published reports that indicate professionalism erodes as aspiring physicians proceed higher in medical training.[4] The main limitation of our study is that all plagiarism, including semantic plagiarism[3] and ghostwriting,[5] could not be detected by the software.

Corresponding Author: *Linda S. Nield, MD; Department of Medical EducationWest Virginia University School of Medicine, Morgantown WV, USA (e-mail: lnield@hsc.wvu.edu).*

REFERENCES

1. **Segal S**, **Gelfand BJ**, **Hurwitz S**, **Berkowitz L**, **Ashley SW**, **Nadel ES**, **Katz JT**. Plagiarism in residency application essays. *Ann Intern Med.* 2010;153(2):112–20.
2. **Parks LJ, Sizemore DC, Johnstone RE**. Plagiarism in Personal Statements of Anesthesiology Residency Applicants. *A A Case Rep.* 2016;6(4):103–5.
3. **Walchuk K.** An Examination of the efficacy of the plagiarism detection software program *Turnitin.* Thesis for the Degree of Master of Education. Graduate Department of Education in the University of Ontario Institute of Technology. Copyright 2016. Available at: https://ir.library.dc-uoit.ca/bitstream/10155/717/1/Walchuk_Kevin.pdf. Accessed on April 17, 2017.
4. **Mostaghimi A**, **Olszewski AE**, **Bell SK**, **Roberts DH**, **Crotty BH**. Erosion of digital professionalism during medical students' core clinical clerkships. *JMIR MED Educ.* 2017;3(1):e9. https://doi.org/10.2196/mededu.6879.
5. **Knapp JC**, **Hulbert AM**. Academic Dishonesty. In:Ghostwriting and the Ethics of Authenticity. New York: Palgrave Macmillan; 2017:85–105.