Assessing Academic Integrity Using SafeAssign
Plagiarism Detection Software

David C. Ison, Ph.D.
Kimberly J. Szathmary, Ph.D.
Embry-Riddle Aeronautical University – Worldwide

Abstract

Higher education has struggled with the problem of plagiarism since institutions were first founded. In the current educational environment where technology and access are integral to institutional functionality and student learning, the concern about plagiarism has been elevated with claims that the availability of sources, commonly sought from the Internet, has made the incidence plagiarism worse. In light of this sentiment and research literature that indicates a systemic problem, it is critical that institutions are aware of their current academic integrity environment. Such information is necessary to address problems, if they exist, and to best manage integrity issues as well as how to assist students in avoiding these quandaries. This study sought to provide baseline data for a large, online, aviation oriented postsecondary institution. A sample of 659 student records were collected in this study of which 520 (78.9%) included a SafeAssign index. The uncorrected SafeAssign index mean was 22.2% ($Md = 17\%$, $s^2 = 390.9$, $s = 19.8$). Among the uncorrected SafeAssign submissions, 216 (41.5%) had initial report values in excess of 15% similarity and 57 (11.0%) had originality indices in excess of 40%. When making corrections for potential false positive results, the mean corrected SafeAssign index was 20.9% ($Md = 17\%$, $s^2 = 220.4$, $s = 14.8$) with 206 (39.6%) of papers equaling or exceeding the 15% threshold and 19 (3.7%) of corrected cases had originality indices in excess of 40%. Additional statistical analysis was conducted to evaluate the relationship between SafeAssign values and final grades. Also, comparisons were made with previous studies. The findings indicate that the institution needs to evaluate its plagiarism policies and student education on the topic. It also highlights the importance of such studies at other institutions to determine benchmarks. Suggestions for future research are provided.

While the problem of plagiarism has existed since the beginning of the sharing of ideas in text and art, it is purported in exigent research that the incidence of plagiarism is widespread and has been rapidly growing in conjunction with the increasing prevalence of the Internet (Bretag, 2013; Evering & Moorman, 2012; Jones, 2011). Of particular concern is that the online research environment makes plagiarism extremely tempting and easy to conduct from the simple cutting and pasting of materials to the ability of students to purchase entire papers for submission from so-called term paper mills (Embleton & Helfer, 2007). Plagiarism occurs at all levels of education from K-12 through the doctorate. Several studies have highlighted the incidence of plagiarism at the culminating stages of graduate degrees within theses and dissertations (Grose, 2006; Ison, 2012, 2014, 2015; Jones, 2011). These findings have also surfaced within the media with high profile individuals having their doctorates revoked as well as cases where numerous students have had their degrees rescinded due to plagiarized theses and dissertations (Ison, 2015). Yet aside from these limited high profile cases, the majority of evidence concerning student plagiarism is a result of self-report surveys of anecdotal evidence rather than empirical data. Thus the knowledge about actual prevalence of plagiarism in higher education, as well as in general, is rather

limited (Ison, 2015; Walker, 2010). As aviation programs have progressed towards a more competency based approach to assessment, especially in terms of how they are evaluated for program-specific accreditation, writing assignments become ever more prevalent and critical to curricula. Moreover, as aviation higher education has migrated to a more research focus, the issue of academic integrity is central to perceptions of veracity and quality (Southeast Oklahoma State University, 2008).

More concerning is that many institutions have little, if any, awareness of the current levels of the occurrence of plagiarism and related misconduct on campus (Novotney, 2011). Some faculty and students may not even understand their responsibilities in terms of academic integrity (Mahmud & Bretag, 2013). As noted by Bretag (2013), Brimble, and Stevenson-Clarke (2005), as well as Cole and Swartz (2013), it is paramount that colleges and universities grasp the level of misconduct residing within such institutions, particularly in light of the concerns about the negative influence of the Internet coupled with the rapid growth of online education. Therefore, it has been advocated that higher education stakeholders determine baseline data concerning the current state of academic integrity within their institutions. In particular, milestone documents are of particular concern (e.g., theses, dissertations, and similar rigorous works) as they are critical components of a course of study (Bretag, 2013; Chertok, Barnes, & Gilleland, 2014; Cole & Swartz, 2013; Ison, 2012, 2015; Novotney, 2011). This study sought to establish the existing levels and severity of plagiarism, and faculty handling thereof, in capstone courses at a large, online aviation related institution.

Purpose

The purpose of this study was to ascertain the current incidence and severity of plagiarism within graduate capstone courses at a large, online aviation related institution as determined by SafeAssign originality reports. The resultant data will provide administration, faculty, and other stakeholders with empirical evidence of the academic integrity environment at the study institution to determine the adequacy and effectiveness of current policies and procedures in addition to being able to monitor trends in integrity over time. This study will also provide the methodological basis for future studies on integrity in other parts and levels of study at the study institution and others.

Literature Review

Defining Plagiarism

The American Psychological Association (2010) defines plagiarism as instances where authors "present the work of another as if it were their own work" (p. 16). A plethora of literature addresses the prevalence of various types of academic misconduct amongst students (McCabe, 2009; Scanlon, 2003; Walker, 2010). College students are generally exposed to definitions of plagiarism and the penalties that will be imposed should they be found guilty of committing plagiarism (Tabor, 2013). Course catalogs, syllabi, and course websites define plagiarism and often link to various mediums where the topic is further explained. Certainly, by the time they reach their final capstone course, students have been thoroughly exposed to the issue, and accurate citations and references have been practiced in a host of course term papers and other assignments.

Plagiarism is sometimes labeled as either intentional or unintentional plagiarism. When committing intentional plagiarism, the author intentionally attempts to mislead the reader. Such an offense might be accomplished by purchasing a paper from a paper-writing website, using creative methods to avoid detection by plagiarism detection software, or by portraying material as paraphrased and citing the source incorrectly. Alternatively, unintentional plagiarism most often occurs when students are not instructed in proper paraphrasing

and citation (Gilmore, Strickland, Timmerman, Maher, & Feldon, 2010). It is often difficult to determine whether plagiarism is intentional or unintentional without questioning the student (Gilmore et al., 2010).

A subset of plagiarism is self-plagiarism, in which authors recycle portions of their own previous work. In academic circles, reuse of one's previously submitted or published work is generally regarded as acceptable if cited appropriately, though in limited amounts (Bretag & Mahmud, 2009). The lack of clear guidelines leaves academics to rely largely on the concept of *fair use*, which is itself a gray area. In any case, appropriate attribution is required.

Previous Studies on Plagiarism in Higher Education

While subjects of academic integrity, including plagiarism, have been an assumed warrant constant attention within higher education since its advent, it has more frequently come into the spotlight in research literature due to the concern that the Internet and online education have had a negative influence on the moral environments at colleges and universities across the globe (Brimble & Stevenson-Clarke, 2005; Ison, 2012; McCabe, 2009; Walker, 2010). Concurrently, there has been increasing interest by researchers and related stakeholders to go beyond simply collecting student opinions about academic integrity and plagiarism instead quantifying the actual incidence of such misconduct in various educational environments. This approach provides a clearer, empirically-based depiction for faculty and administrators as to what is actually occurring within their institutions. Such feedback is necessary to understand if, and to what extent, there may be an integrity problem that may need to be addressed (Cabral-Cardoso, 2004; Walker, 2010). Thus this review focuses primarily on quantitative studies utilizing direct observations of plagiarism pervasiveness.

Examining two major assignments in sophomore year courses at a postsecondary institution in New Zealand, Walker (2010) found 25% of what was submitted ($n = 569$) exhibited indications of plagiarism. A study of roughly 1,000 students' work which was submitted without knowledge that it would be computer analyzed for originality had 16.5% textual overlaps with source material. Within the same study, an equivalently sized sample of students who knew their work would be scrutinized by plagiarism detection software had a mean overlap of 10% (Heckler, Rice, & Bryan, 2013). Martin et al. (2009) utilized originality checking software to analyze submissions within postsecondary business courses and found that 61% of the sample ($n = 158$) showed evidence of plagiarism with these works having a mean similarity of 11%. A subsequent and much larger study ($n = 40,000$) using detection software across a broader range of institutions and subject areas found that over 30% of submissions had evidence of plagiarism (Martin et al., 2011). Exploring graduate level work, Chao, Wilhelm, and Neureuther (2009) examined 116 papers written by master's in business administration students. Among those works, 39% were identified as to contain suspected plagiarized material. In another study of master's student work, Homberg and McCullough (2015) ($n = 68$) discovered that 67.8% of theses had highly suspicious levels of textual overlap with source material. A study by Ison (2015) discovered that among 184 dissertations written in recent years (since 2010), 52% had suspicious levels of commonality with source material.

Not only is it clear that students are actively engaging in plagiarism at disquieting rates, the nature of such occurrences are also concerning. In a study of work by 700 students, Scanlon and Neumann (2002) found that 25% included materials that was literally "cut and pasted" from the sources. Walker (2010) also found that among papers in their sample, 14% had verbatim phrases and sentences directly drawn from source material although students made an attempt, albeit inaccurate, to cite such material while another 11% had the exact text as found in source material but without any citation. Gilmore et al. (2010) discovered "instances of cutting and pasting large chunks of text without quotation or citation" (p. 18).

Another notable commonality within recent related literature is the concern that the online environment is fostering academic dishonesty. This is of particular concern as online education has become more ubiquitous. Kennedy, Nowak, Raghuraman, Thomas, and Davis (2000) stated that students felt as though it was simpler to cheat in the virtual environment and, as such, as online education grows, the authors posited that the occurrence of academic misconduct such as plagiarism will also escalate. Lanier (2006) discovered "that cheating was much more prevalent in online classes compared to traditional lecture courses" (p. 244). Moreover, the Internet has been reported as an easy pathway for students to copy work as research has transformed, in many cases, into an exercise of cutting and pasting of source material (Auer & Krupar, 2001). Postle (2009) found that there has been a clear increase in plagiarism with significant inappropriate use of text increasingly borrowed from online sources. Reinforcing this premise, Townley and Parsell (2004) stated that the Internet allows students to more efficiently misappropriate materials from both other students and the literature. In a 2005 study, McCabe found that 60% of the 80,000 graduate and undergraduate papers analyzed from 83 institutions included cut-and-pasted material from the Internet. Similarly, Selwyn (2008) discovered that approximately 60% of students reported inserting verbatim text from online sources into their work.

As students progress to higher levels of education, such indiscretions have even greater implications. Master's and doctoral degrees hang greater weight on significant written works (e.g., theses and dissertations). If these efforts are marred with plagiarized material, they call into question the merit of the awarded degree. As cases of plagiarism become rampant, such can have negative consequences on academic reputation and accreditation Other recent studies have highlighted actual cases which have resulted in academic scandals that eventually led to significant negative media attention. Further, in some of these instances, institutions revoked the graduate degrees of the transgressors (Powers, 2008; Riog, 2010).

One particular problem revealed within the literature is that institutions often do not have a real sense of the prevalence of plagiarism within their classes (Brimble & Stevenson-Clarke, 2005; Ison, 2012; 2015; Novotney, 2011; Powers, 2008; Roig, 2010). While many colleges and universities have dedicated personnel or entities charged with monitoring and supervising academic integrity, some do not (Brimble & Stevenson-Clarke, 2005; Embry-Riddle Aeronautical University, 2015; Grand Canyon University, 2015; Ison, 2015). The study institution appears to fit what the International Center for Academic Integrity (2012) classifies as "Stage One: Primitive" in terms of its awareness and handling of plagiarism: "this stage describes a school with no policy or procedures (or minimalist ones) and where there is great variation in faculty and administrative handling of cheating" (para. 1). While the institution has academic integrity statements and procedures, there is no one central database or tracking thereof, no office of academic integrity, nor a central figure overseeing the process across the campus. These realities reinforce the need to appraise the current academic integrity environment at the study institution.

Method

Sample characteristics

Research participants were graduate-level students seeking a master's degree at a large, online aviation related institution. All students were enrolled in the institution's capstone course, which has, as its prerequisite, the completion of at least three graduate-level courses prior to enrollment. Admission to the master's degree programs require students to have achieved at least a bachelor's degree.

Sampling

All 2014 capstone students whose SafeAssign scores and final grade point values were available were included in the sample. Students were excluded if either the SafeAssign score was not present, or the final grade point value awarded was not available in the grading management system.

Sample size

Among 659 initial student data captures, only 456 students' records that included both the SafeAssign score and the final grade point value were utilized. This far exceeded the necessary sample sizes for the statistical testing utilized in this study at the .80 level per G*Power sample size and power software calculations.

Measures

A total of 38 individual course sections were evaluated to determine if the capstone papers associated with students enrolled in those sections met the criteria. Some instructors managed the capstone submission and evaluation process independent of the learning management system, and either SafeAssign scores or final grade point values were not available to the researchers in the course grade center. The researchers retrieved records by downloading the data from the final course grade centers into MS Excel spreadsheets, and by individually retrieving and adding SafeAssign scores to those spreadsheets for each student record, where available.

Research Design

This descriptive study sought to define the pervasiveness of plagiarism in the graduate capstone submissions at a large, online aviation related institution. All 2014 graduate capstone courses were mined for SafeAssign indices and final grade point values. Data was recorded for each student whose data included both a SafeAssign index and a final grade point value. Where individual student SafeAssign indices exceeded 15%, reports were examined by the researchers to determine erroneous indications, and indices were corrected manually. This 15% threshold is outlined by SafeAssign as the cutoff for suspect plagiarism while those above 40% have high evidence (Blackboard, 2016). Corrected errors included quoted materials that were counted towards the total percentage, typical academic writing and statements (e.g. related to hypothesis testing or the statement of hypotheses), and copyright or document identification language. Although SafeAssign generally can successfully identify quoted material, there are, on occasion, some legitimate overlaps which are flagged by the software. This type of quality checking has been advocated by similar studies by Batane (2010) and Bretag (2013). Where no error was indicated, the indices were not adjusted.

Results

Missing Data

Because of the lack of reporting by faculty teaching and students enrolled in the capstone courses, there were some missing values in some cases among grades and SafeAssign indices. So as not to corrupt statistical calculations, these cases were omitted from calculations individually (descriptive statistics) or pair-wise (correlation analysis).

Descriptive Statistics

Among the 659 student records collected in this study, 486 (73.7%) included a point value (grade) for the capstone submission and 520 (78.9%) included a SafeAssign index. Only 456 (69.2%) records included both a grade value and a SafeAssign index. The mean grade for the sampled capstones was 91.3% ($Md = 92\%$, $s^2 =$ 2323.5, $s = 48.2$). See figure 1 for the histogram of reported grades. The uncorrected SafeAssign index mean was 22.2% ($Md = 17\%$, $s^2 = 390.9$, $s = 19.8$). See figure 2 for the histogram of the uncorrected SafeAssign indices. Among the uncorrected SafeAssign submissions, 216 (41.5%) had initial report values in excess of 15% similarity and 57 (11.0%) had originality indices in excess of 40% (value at which SafeAssign deems to have a high likelihood of significant plagiarism).

Cases with index values in excess of 15% were manually examined to remove any erroneous indication of overlap (false positives). In some cases, no correction was warranted. The mean correction applied to the SafeAssign indices was 6.6%. The resultant mean corrected SafeAssign index was 20.9% ($Md = 17\%$, $s^2 = 220.4$, $s = 14.8$) with 206 (39.6%) of papers equaling or exceeding the 15% threshold and 19 (3.7%) of corrected cases had originality indices in excess of 40%. See figure 3 for the histogram of corrected SafeAssign indices.
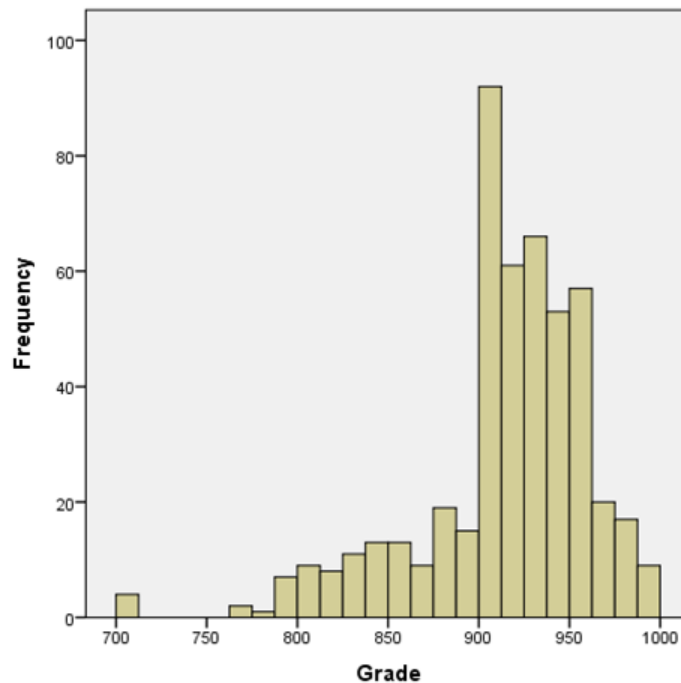

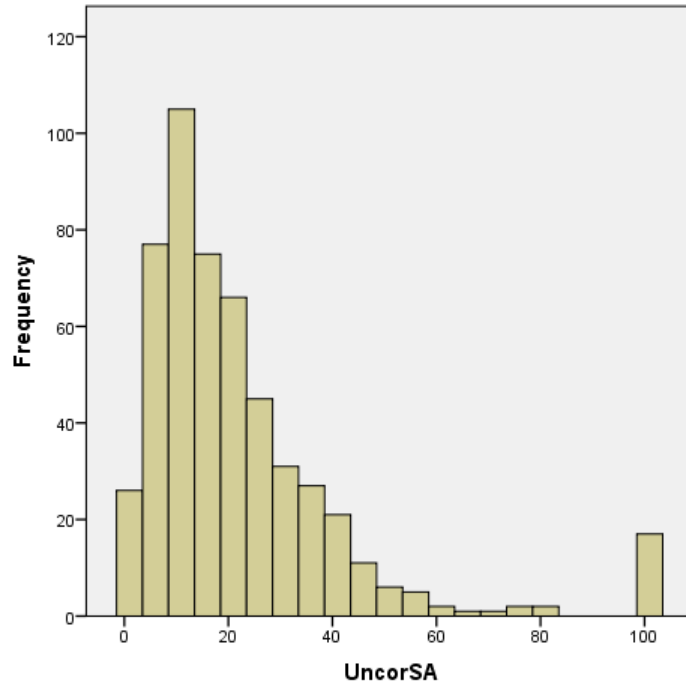*Figure 1*. Frequency of grade values of sampled capstones

Figure 2. Frequencies of uncorrected SafeAssign index values among sampled capstones
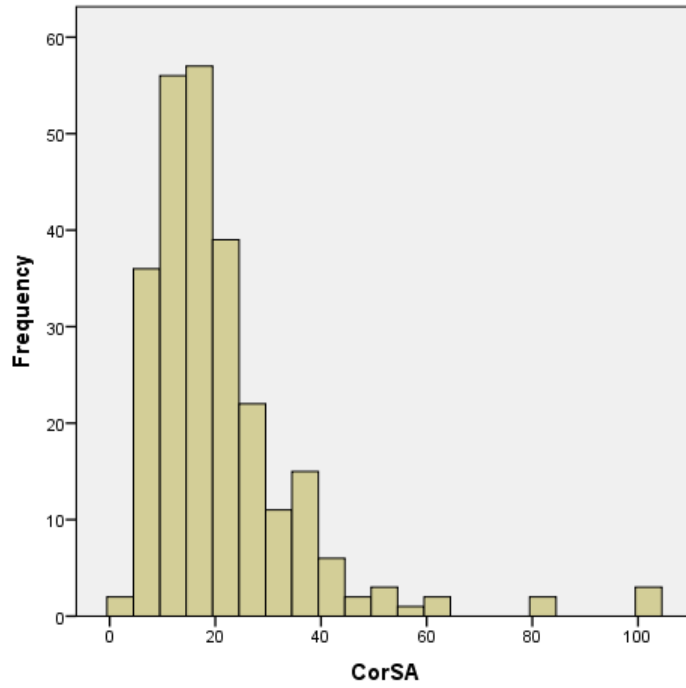

Figure 3. Frequencies of corrected SafeAssign index values among sampled capstones

Correlation Analysis

Because the grade and SafeAssign index data were not normally distributed, relationships between data sets were evaluated with the non-parametric correlation Kendall's tau-b ($r_{\tau b}$). The selection of this test was guided by literature by Howell (2013) in which the case was made that Kendall's tau-b is generally superior to Spearman's rho and takes into account cases of pairwise ties. Correlations could only be calculated among available pairs of data.

Results of the analysis indicated a negligible negative correlation, albeit non-significant, relationship between grade and the uncorrected index values, $r_{\tau b} = -0.041$, $p = 0.198$, $d = 0.129$.[1] Similarly, there was effectively no relationship, yet again negative and non-significant, between grade and corrected index values, $r_{\tau b} = -0.012$, $p = 0.785$, $d = 0.038$. As noted by Lewis-Beck, Bryman, and Liao (2004), findings such as these when sample sizes are large, as is true in this case, indicate that no noteworthy correlation exists.

Comparisons with Previous Studies

Although not a central premise of the current study, comparisons of findings with other related studies places the aforementioned data in perspective of the exigent literature. Table 1 compares the current findings with those in recent studies. Specifically, the percentage of papers in each study that were suspected to contain plagiarism are provided. Table 2 provides the mean similarity indices of the current study in comparison to those in other studies.

Table 1.
Comparison of results: Percentage of papers with suspected plagiarism.

Study	Documents with plagiarism (%)
Ward (2002)	25
Walker (2010)	25
Martin et al. (2011)	30
Gilmore et al. (2010)	38
Chao et al. (2009)	39
Current Study	39.6 (corrected); 41.5 (uncorrected)
Ison (2012)	60
Martin et al. (2009)	61
Holmberg and McCullough (2015)	67.8
Batane (2010)	100

When comparing the corrected indices of the current study with those given in Ison (2015) and Ison (2012), a Kruskal-Wallis *H* test showed that there was a statistically significant difference among the findings of the studies, $\chi^2(2) = 43.27$, $p < 0.001$, with a mean rank index of 550.78 for the current study, 424.57 for Ison (2015), and 444.01 for Ison (2012). *Post-hoc* Mann-Whitney *U* tests were conducted on pairs which indicated that there were differences between the indices in the current study and both Ison studies (2015: $U = 56524$, $p < 0.001$, $r = 0.17$; 2012: $U = 35890$, $p < 0.001$, $r = 0.21$), but no differences existed between the Ison studies (2015 & 2012: $U = 24406$, $p = 0.322$, $r = 0.05$).

---

[1] Effect sizes calculated via $r = \sin(.5\pi\tau)$ then $d = 2r/[(1-r^2)^{.5}]$ per Walker (2003).

Table 2

Comparison of results: Similarity/Originality indices

Study	Similarity/Originality Indices (%)
Martin et al. (2011)	10.6
Ison (2012)	15.1
Ison (2015)	13.4
Batane (2010)	20.5
Current Study	20.9 (corrected); 22.2 (uncorrected)

Additional comparisons were made with studies that specifically focused on master's capstone/theses utilizing Chi square tests for independence with Yates' correction. When comparing counts of corrected capstones that had suspicious levels of textual overlap (>15%) with theses indicating evidence of plagiarism in Holmberg and McCullough (2015), a significant relationship was present (χ^2 [1, N = 708] = 29.341, p < 0.001). When comparing study data with that in Martin et al. (2009), a significant relationship was present (χ^2 [1, N = 784] = 28.826, p < 0.001). In both of the aforementioned cases, the samples in Holmberg and McCullough (2015) as well as Martin et al. (2009) had higher levels of suspected plagiarism than that of this study. Yet in a comparison between the present study data and that of Chao et al. (2009), no significant relationship was present (χ^2 [1, N = 732] = 1.105, p = 0.313).

Discussion

The results of this study should cause great concern not only to the study institution, but also to academic researchers in general. The findings of this study reinforce the sentiment that academic integrity issues are a factor for a wide range of institutions and subjects of study. With more than 39% of graduate student capstones containing at least some plagiarized material, standards for research are not being taught and enforced at a level to ensure ethical behavior. Furthermore, the final grade point values are not positively correlated to the high safe assign indices, indicating that the capstone instructors do not always monitor the SafeAssign reports and enforce academic integrity expectations.

Of note in this study, the institution does not enforce a limiting score for SafeAssign indices, and grading and enforcement are at the instructor's discretion. Faculty members are expected to investigate scores where the possibility of plagiarism exists, but specific training is not provided. Assistance in interpreting the information is available, but not often sought.

It is critical to recognize that some of the studies used for comparison utilized samples of different types of educational milestones (e.g. papers, theses, and dissertations) at different levels (undergraduate versus graduate). Therefore, some of the differences in similarity indices may be explainable due to this variation among sample constructs. What is clear, regardless of sample, is that plagiarism is likely occurring in all facets of education and warrants the attention of stakeholders. Without similar such studies, institutions are not able to identify, monitor, or correct these deficiencies.

Conclusion

This study further cements the findings of a large body of research that indicates the significance of plagiarism in postsecondary education. While the percentage of capstone papers with plagiarism in this study is much lower than percentages in many other studies (Batane, 2010; Holmberg & McCullough, 2015; Ison, 2012; Martin et al., 2009), the percentage was still much higher than others. Lack of monitoring of SafeAssign indices and the awarding of passing grades to students whose work contains probable plagiarism could be construed as legitimizing misconduct, and could encourage new faculty to overlook such behavior, versus correcting and promoting ethical behavior.

Recommendations

The results of this study should motivate institutional stakeholders to address the problem by increasing faculty awareness and training in the area of plagiarism. Students should also receive at least minimal training in research, paraphrasing, citing, and referencing early in the graduate program, and should be monitored and corrected throughout their journey to the capstone course. While establishing a standard for formatting, references, and citation is a first step, ensuring faculty have the tools to correct students' errors in paraphrasing, citations, and ethics in research is imperative. While this study provides a baseline for the incidence of plagiarism in capstones, the institution must move forward by establishing expectations for faculty that include actions to take for submissions with high similarity indices as well as criteria for grading that requires the faculty member to delve deeper into a capstone submission with a high SafeAssign index in an effort to rule out potential plagiarism. Faculty members carry some measure of responsibility to the institution to ensure that instances of plagiarism are addressed. Hopefully, this study will lead to positive change in the institution's capstone course product. As the institution moves to forward in technological advances and changes the time to set expectations, policies, and procedures is now.

References

About SafeAssign. (2010). *Blackboard Inc.* Retrieved from http://library.blackboard.com /ref/a86c3648-80a5-43cc-8fed-b3f5d24518ce/Content/SafeAssign /About%20SafeAssign.htm

American Psychological Association. (2010). *Publication manual of the American Psychological Association*, (6th ed.). Washington, DC: Author.

Auer, N. J., & Kupar, E. M. (2001). Mouse click plagiarism: The role of technology in plagiarism and the librarian's role in combating it. *Library Trends, 49*(3), 415-432.

Batane, T. (2010). Turning to Turnitin to fight plagiarism among university students. *Journal of Educational Technology & Society*, *13*(2), 1-12.

Blackboard. (2016). *SafeAssign originality report*. Retrieved from https://en-us.help.blackboard.com/Learn/9.1_2014_04/Instructor/100_Assignments/025_Use_Safe Assign/010_SafeAssign_Originality_Reports

Bretag, T. (2013). Challenges in addressing plagiarism in education. *PLoS Med, 10*(12), e1001574. doi:10.1371/journal.pmed.1001574

Bretag, T., & Mahmud, S. (2009). Self-plagiarism or appropriate textual re-use? *Journal of Academic Ethics, 7,* 193-205. doi: 10.1007/s10805-009-9092-1

Brimble, M., & Stevenson-Clarke, P. (2005). Perceptions of the prevalence and seriousness of academic dishonesty in Australian universities. *The Australian Educational Researcher, 32*(3), 19-44.

Cabral-Cardoso, C. (2004). Ethical misconduct in the business school: A case of plagiarism that turned bitter. *Journal of Business Ethics, 49*, 75-89.

Chao, C. A., Wilhelm, W. J., & Neureuther, B. D. (2009). A study of electronic detection and pedagogical approaches for reducing plagiarism. *The Journal of Research in Business Education*, *51*(1), 31.

Chertok, I. R., Barnes, E. R., & Gilleland, D. (2014). Academic integrity in the online learning environment for health sciences students. *Nurse Education Today, 34*(10), 1324-1329. doi:10.1016/j.nedt.2013.06.002

Cole, M. T., & Swartz, L. B. (2013). Understanding academic integrity in the online learning environment: A survey of graduate and undergraduate business students. In *Proceedings of the 20th Annual ASBBS Conference* (pp. 738-746).

Embleton, K., & Helfer, D. S. (2007). The plague of plagiarism and academic dishonesty. *Searcher, 15*(6), 23.

Evering, L. C., & Moorman, G. (2012). Rethinking plagiarism in the digital age. *Journal of Adolescent & Adult Literacy, 56*(1), 35-44. doi:10.1002/JAAL.00100

Gilmore, J., Strickland, D., Timmerman, B., Maher, M., & Feldon, D. (2010). Weeds in the flower garden: An exploration of plagiarism in graduate students' research proposals and its connection to enculturation, ESL, and contextual factors. *International Journal for Educational Integrity*, *6*(1).

Grand Canyon University. (2015). *Academic integrity*. Retrieved from http://www.gcu.edu/Academics/Academic-Integrity.php

Grose, T. (2006). The burden of plagiarism: The cheating scandal at Ohio University has raised all sorts of questions. *ASEE (American Society for Engineering Education) Prism, 16*(3). Retrieved from http://www.prism-magazine.org/nov06/feature_plagiarism.cfm

Heckler, N. C., Rice, M., & Bryan, C. H. (2013). Turnitin systems: A deterrent to plagiarism in college classrooms. *Journal of Research on Technology in Education, 45*(3), 229-248.

Holmberg, M., & McCullough, M. (2006). Plagiarism in science and technology master's theses: A follow-up study. *New Review of Information Networking, 12*(1-2), 41-45. doi: 10.1080/13614570601136255

Howell, D. C. (2013). *Statistical methods for psychology* (8th ed.). Belmont, CA: Wadsworth.

International Center for Academic Integrity. (2012). *Resources*. Retrieved from http://www.academicintegrity.org/icai/resources-4.php

Ison, D. C. (2012). Plagiarism among dissertations: Prevalence at online institutions. *Journal of Academic Ethics, 10*(3), 227-236. doi:10.1007/s10805-012-9165-4

Ison, D. C. (2014). Does the online environment promote plagiarism: A comparative study of dissertation from brick-and-mortar versus online institutions. *MERLOT Journal of Online Learning and Teaching, 10*(2), 272-281.

Ison, D. C. (2015). The influence of the Internet on plagiarism among doctoral dissertations: An empirical study. *Journal of Academic Ethics*, *13*(2), 151-166.

Jones, D. L. (2011). Academic dishonesty: Are more students cheating? *Business Communication Quarterly, 74*(2), 141-150. doi:10.1177/1080569911404059

Lewis-Beck, M. S., Bryman, A., & Liao, T. F. (Eds.). (2004). *Encyclopedia of social science research methods.* (Vols. 1-3). Thousand Oaks, CA: Sage Publications, Inc. doi: http://dx.doi.org/10.4135/9781412950589

Mahmud, S., & Bretag, T. (2013). Postgraduate research students and academic integrity: "It's about good research training." *Journal of Higher Education Policy and Management, 35*(4), 432-443.

Martin, D. E., Rao, A., & Sloan, L. R. (2009). Plagiarism, integrity, and workplace deviance: A criterion study. *Ethics & Behavior*, *19*(1), 36-50.

Martin, D., Rao, A., & Sloan, L. (2011). Ethnicity, acculturation, and plagiarism: A criterion study of unethical academic conduct. *Human Organization*, *70*(1), 88-96.

McCabe, D. L. (2005). Cheating among college and university students: A North American perspective. *International Journal for Educational Integrity, 1*(1).

McCabe, D. L. (2009). Academic dishonesty in nursing schools: An empirical investigation. *Journal of Nursing Education, 48*(11), 614-623.

Novotney, A. (2011). Beat the cheat. *American Psychological Association Monitor, 42*(6), 54.

Postle, K. (2009). Detecting and deterring plagiarism in social work students: Implications for learning for practice. *Social Work Education, 28*(4), 351-362. doi:10.1080/02615470802245926

Powers, E. (2008, February 22). "Distinguished" no longer. *Inside Higher Ed*. Retrieved from http://www.insidehighered.com/news/2008/02/22/ohio

Roig, M. (2010, June). Plagiarism: An ounce of prevention. *ASQ Higher Education Brief*. Retrieved from http://www.asq.org/edu/2010/06/continuous-improvement/plagiarism-an-ounce-of-prevention-.pdf

Scanlon, P. M. (2003). Student online plagiarism: How do we respond? *College Teaching, 51*(4), 161-165.

Scanlon, P. M., & Neumann, D. R. (2002). Internet plagiarism among college students. *Journal of College Student Development, 43*(3), 374-385.

Selwyn, N. (2008). "Not necessarily a bad thing ...": A study of online plagiarism amongst undergraduate students. *Assessment & Evaluation in Higher Education, 33*(5), 465-479.doi:10.1080/02602930701563104

Southeastern Oklahoma State University. (2008). *Outline for a self-study report: Form 204*. Retrieved from http://www.se.edu/dept/academic-affairs/files/program-reviews-special-accreditations/docs/SE%20AABI%20Self%20Study%202011.pdf

Southern Association of Colleges and Schools. (2014). Policies and publications. Retrieved from http://www.sacscoc.org/policies.asp

Tabor, E. L. (2013). *Is cheating always intentional? The perception of college students toward the issues of plagiarism* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses. (Order No. 3557437)

Townley, C., & Parsell, M. (2004). Technology and academic virtue: Student plagiarism through the looking glass. *Ethics and Information Technology, 6*(4), 271-277. doi:10.1007/s10676-005-5606-8

Walker, D. C. (2013). Converting Kendall's Tau for correlational or meta-analytic analyses. *Journal of Modern Applied Statistical Methods, 2*(2), 525-530.

Walker, J. (2010). Measuring plagiarism: Researching what students do, not what they say they do. *Studies in Higher Education, 35*(1), 41–59.

Ward, J. (2002, May 23). University sees rise in reports of cheating. *Washington Times*, p. B-1.

Do NTSB Statistics Support Current FAA Third Class Medical Policy?

Don Morris
Southern Illinois University, Carbondale

Abstract

Third Class Medical Reform is currently being considered by the FAA and by Congress. This study analyzes past incident data gathered by the NTSB from 1982 to the present to determine if a link exists between the amount of time that has elapsed since a pilot's medical certification and his or her likelihood of involvement in an incident or accident caused by medical incapacitation or impairment. A control population of 26,987 incidents was formed from all recorded incidents in which the pilot held a valid third class medical. This population was used to compare with a smaller population of 172 incidents for which the NTSB has determined that medical incapacitation or impairment not related to illegal drug or alcohol use was at least a factor. A Chi-Squared distribution returned a 17% probability that the two populations were identical. Incidentally, descriptive statistics of the control population show an unexpected negative linkage between recentness of medical certification and likelihood of being involved in an incident ($R^2 = .821$). This counterintuitive finding indicates that recent third class medical certification correlates to increased pilot risk, rather than increased levels of safety.

Background

If a pilot is involved in an incident or accident, they are required to notify the National Transportation Safety Board (NTSB) (49 CFR 830.15). The NTSB investigates the incident and, whenever possible, determines contributing factors. They record these along with details of the incident in the enhanced Accident Data Management System (eADMS) database. The database has hundreds of thousands of entries, and is publicly available. One of the details recorded for each occurrence in the database is the date of the pilot's last medical exam.

In order to maintain pilot safety, the Federal Aviation Administration (FAA) requires pilots to see a designated medical examiner (DME) to certify that they are medically fit to fly (14 CFR 61.23). The DME checks the pilot's physical condition, current medications, and medical history. If the DME finds that the pilot is capable of safely operating an aircraft and is unlikely to develop any conditions that would put them in danger while operating an aircraft, they are certified as medically fit. Unfortunately, as time elapses, a pilot's condition may unexpectedly deteriorate. This is why medical certificates only last a certain period of time before they must be renewed. Current FAA regulations require medical certificates to be renewed in intervals from six to 60 calendar months, depending on the class of medical certificate and the age of the pilot at the time of the exam (Federal Aviation Administration, 2008).

In order to exercise the privileges of an airline transport pilot, a pilot must have a first class medical. Commercial pilot privileges require a second class medical. Private pilot privileges require a third class medical. Currently, third class medical reform is before the FAA and Congress (Aircraft Owners and Pilots Association,

2012; Pilots Bill of Rights 2, 2015). Those advocating reform say that the current system is excessively burdensome and expensive (Aircraft Owners and Pilots Association, 2012; Air Facts Journal, 2013). It is also, they add, redundant. Most pilots have already gone through some form of basic medical fitness check in order to receive a driver's license (Aircraft Owners and Pilots Association …, 2012). They advocate that a State issued driver's license be allowed to replace the FAA specific third class medical certificate.

Proponents of medical reform say there is "almost zero link" between medical certification and safety (Aircraft Owners and Pilots Association, 2012; Goyer, 2014). Opponents of reform point out that past statistics cannot show the incidents and accidents that have been prevented by the system (Webb, 2014), and thus the statistics show that the system is working. However, actual statistical analysis is noticeably absent from the discussion. There seems to be a legitimate reason for this. The researcher's search for published statistical analyses of third class medical effectiveness produced no results. It seems that the actual numbers are not known. This could explain why FAA policy on third class medicals was shaped by a statistical analysis performed on air traffic controllers, and not on private pilots (Aircraft Owners and Pilots Association …, 2012).

It is the researcher's hope that this study will be the beginning of an evidence based discussion on the effectiveness of the third class medical. This study cannot hope to determine whether or not medical reform is wise. It can, however, begin the discussion with a thorough investigation of the historic effectiveness of the third class medical as revealed by past NTSB data.

Rationale

It is logical to assume that any safety enhancement provided by medical certification will be greatest on the day of the medical exam. Over time, as unforeseen changes occur in the pilot's condition, its effect will wane. If the time that has elapsed since the last medical is plotted against the number of incidents, any relationship between these two variables should become apparent. In the absence of unknown codependent factors, a positive correlation between the two factors would indicate that the medical certification program is preventing incidents. If no correlation exists, this would seem to indicate that medical certification program is not preventing incidents. A negative correlation would indicate that a fresh medical certificate is associated with an increased risk of being involved in an incident – which is certainly not the goal of the program. Figure 1 shows the three most likely theoretical distributions based on illustrative data, with the "slight positive correlation" being the normally assumed shape of the graph. The expectation of this correlation is the presumptive basis for current FAA medical policy.
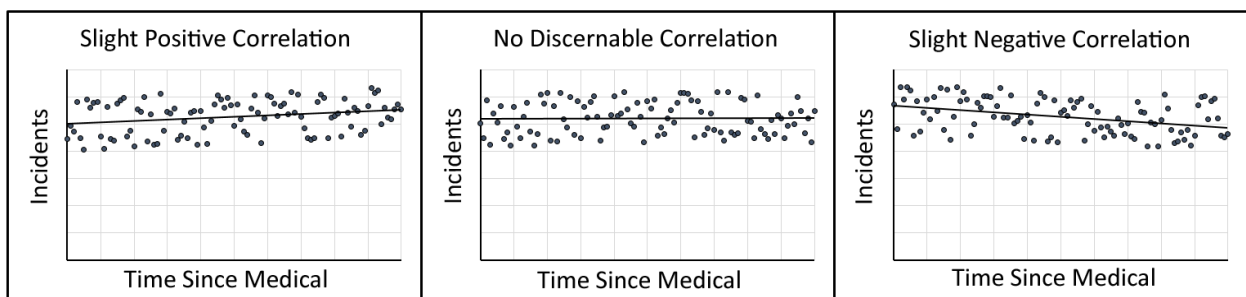

Figure 1. Possible theoretical distributions

Data

This study defines two separate populations both taken from the eADMS database covering the time period from January 1st of 1982 and extending to April 30th of 2015.

1. The first population will be referred to as the control population. It consists of all incidents in which a pilot had a valid third class medical, a valid date of last medical was recorded, and a valid pilot age was recorded. This population includes 26,987 distinct incidents.

2. The second population will be referred to as the research population. It consists of all the incidents out of the control population in which the NTSB determined that impairment or incapacitation (not relating to alcohol or illegal drug usage) was at least a factor. This population includes 172 distinct incidents.

The final data required to complete this study is the percentage of elapsed time between the issuance of the last medical and each incident that has been recorded in the database. Obtaining this data is not simple because the regulations pertaining to medical duration changed twice during the time period covered by the data (Federal Aviation Administration, 2008). Between 1982 and 1996, all third class medicals were valid for 24 calendar months. Between 1996 and 2008, third class medical validity for those under the age of 40 was extended to 36 calendar months. After 2008, third class medical validity for those under 40 years of age was again extended – this time to 60 calendar months. Conditional statements were used to model these regulations, and the percentage of time that had elapsed from the medical at the time of the incident was added to the data. This was done for both populations. Descriptive statistics for each population are shown in table 1.

Table 1.
Time is shown as the percentage of time elapsed from the valid medical time period

Descriptive Statistics from Each Population		
	Control	Research
Sample Size	26987	172
Avg. Time Ellapsed	43.6%	47.2%
Stand. Dev.	28.6%	26.7%

Analysis

Preliminary analysis was the same for both populations. Probability histograms (0 to 100) were constructed by rounding the percentage of time elapsed to the nearest percent, and by doubling the number of occurrences in the 0 and 100 buckets to compensate for the fact that these buckets were only half the width of the rest. Scatter plots were made for each population, and a regression line was calculated for each. These plots are shown in figures 2 and 3, along with regression lines and their respective R^2 values. For simplicity sake, only linear regressions were used in this study, although other unknown models might better fit the data. As can be seen from figure 2, the regression line for the control population has a strongly negative slope. This in counterintuitive, as it appears to associate recentness of medical exam with increasing risk of accident or incident, rather than with increasing levels of safety. Figure 3 shows the regression line for the research

population. It has a similar negative slope, but both the slope and the coefficient of determination are less pronounced.
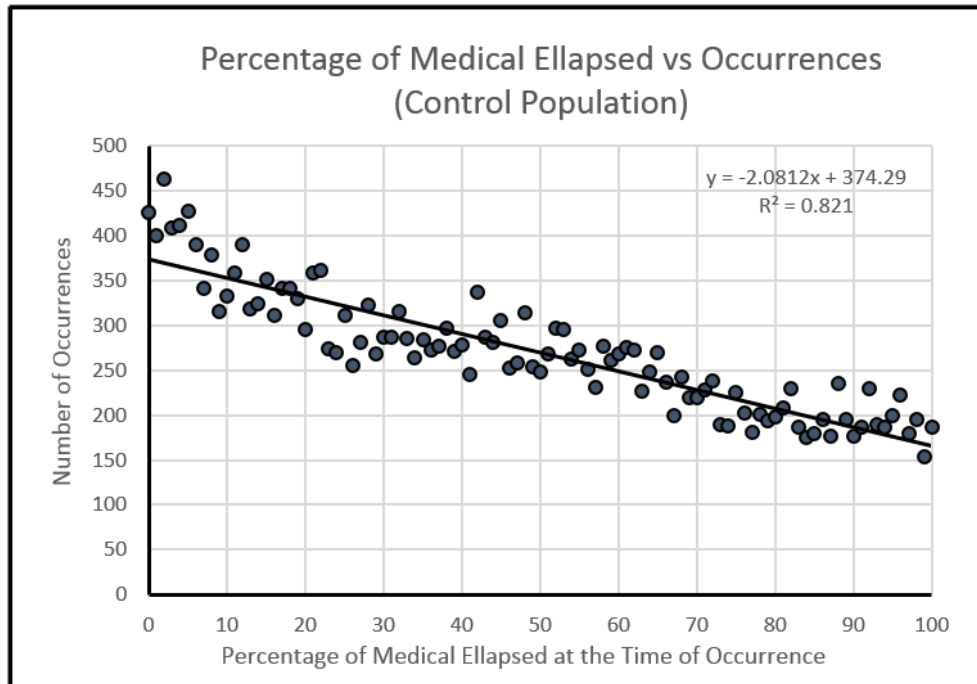

Figure 2. Percentage of medical elapsed vs. occurrences (Control population)
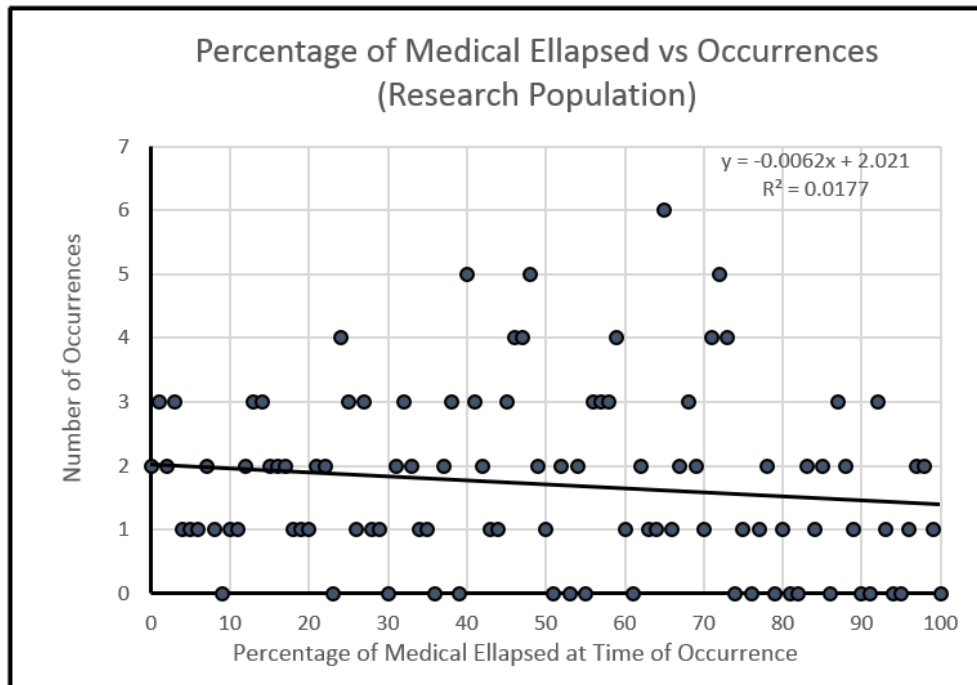

Figure 3. Percentage of medical elapsed vs. occurrences (Research population)

The objective of this study is to determine if a link exists between time elapsed and the risk of being involved in an incident involving incapacitation or impairment. The control population shows a definite link, but this is the likelihood of being in an "incident," and not the likelihood of being involved in an "incident involving impairment or incapacitation." In order to show the latter, it is necessary that the research population be statistically distinguishable from the control population.

In accordance with conventional hypothesis testing logic, two alternative hypotheses can be identified. The research hypothesis is that the research population can be statistically distinguished from the control population. This is to say that an underlying variable other than random chance is statistically unlikely to have created the observed differences between them. The null hypothesis is that the research population cannot be statistically distinguished from the control population. This is to say that the observed differences are within the level that is likely explained by random factors.

In order to test the null hypothesis, the control population was used to calculate how many occurrences would be expected to occur in the research population for each histogram bucket, assuming that the populations were indistinguishable. This was determined by scaling the control population for sample size. The expected occurrences were then compared to the actual occurrences found in the research population using a Chi Squared test for goodness of fit. This test shows a probability of 17% (p=0.17) that the differences between the two populations are due to random factors.
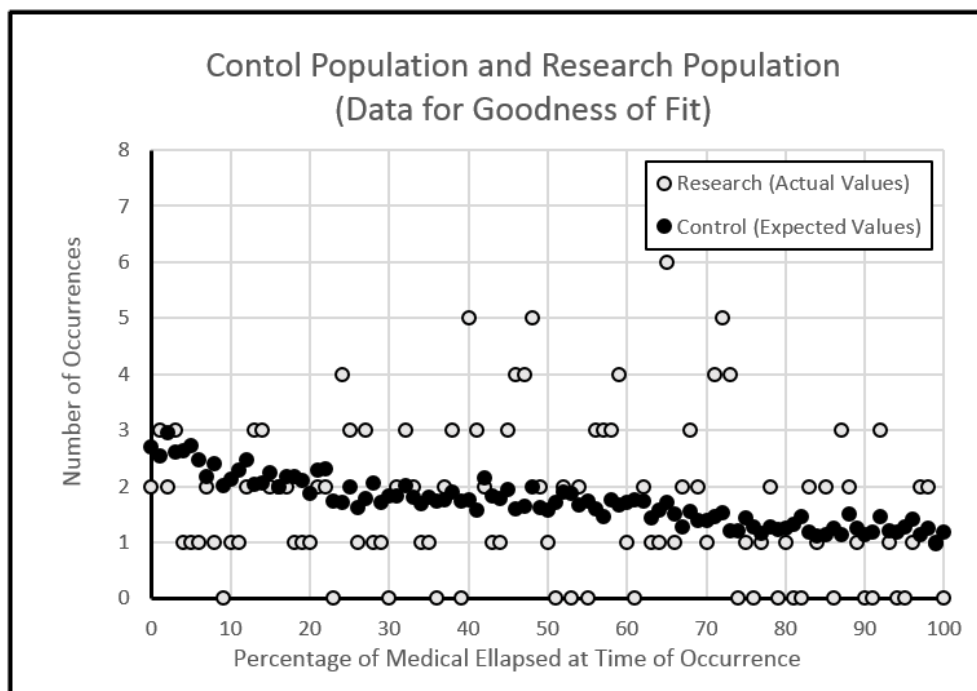

Figure 4. Control population and research population (Data for goodness of fit)

Conclusion

Most studies report a linkage if the probability that the null hypothesis is correct is found to be below 5%. Some studies use a less rigorous standard of 10%. The 17% probability found in this study does not meet

either of these criteria. The statistics do not allow rejection of the null hypothesis, and so do not show a statistically significant linkage between percentage of medical elapsed and likelihood of incident occurrence. However, in the interest of safety, it is important to point out that the data does indicate an 83% likelihood that the two populations are statistically distinguishable in some form, however minute it may be.

While these findings are significant, they pale in significance compared to the unexpected negative correlation between percentage of elapsed medical time and likelihood of incident found in the control population. Statistically speaking, the more time that has elapsed since the third class medical, the less likely a pilot is to be involved in an incident or accident. This result is implied the existence of other unknown codependent variables, as it is not logical to assume that a medical evaluation in itself makes a pilot more dangerous. Logically, there are many possible explanations. The researcher and his colleagues have identified the following:

1. A private pilot who is not intending to fly is unlikely to renew his or her medical until that intention changes. Initial medical certification is certainly tied to an increase in flight activity as training commences. Recent medical recertification is probably tied to a similar increase in flight activity. Even if the per-hour likelihood of being involved in an incident is not increased, the increase in actual hours of flight activity would logically lead to an increased likelihood of incident.
2. Recent medical certification may be tied to complacency about emerging medical conditions that may lead to incapacitation or impairment.
3. Initial medical certification is tied to a lack of experience as a pilot. Recent medical certification is often tied to a lack of recent experience as a pilot. Decreased actual experience and decreased recent experience are both likely to increase the probability of involvement in incidents.

Regardless of the reasons for this statistical correlation, it is very strong. The coefficient of determination (R^2) was 82.1%. This indicates that 82.1% of the observed variation in incident rate can be attributed to the linear regression model. Stated another way, 82.1% of a private pilot's historic risk of being involved in an accident or incident can be determined based solely on how recently he or she was issued his or her last third class medical. The likelihood that this variance could be attributed to random chance was calculated to be nearly zero ($p < 1 \times 10^{-160}$), based on a T test against randomly generated data.

As previously noted, proponents of third class medical reform believe that the current medical certification process is burdensome and expensive. Opponents insist that the process is neither costly nor burdensome (Eidson, 2014; Webb, 2014). It is difficult to imagine why the amount of elapsed time since the issuance of a medical certificate is so strongly correlated to the likelihood of a pilot being involved in an accident or incident if the issuance of a third class medical is trivial to the pilot in question.

References

Air Facts Journal (2013). The great debate: Should the 3rd class medical be eliminated? Retrieved from
http://airfactsjournal.com/2013/03/the-great-debate-should-the-3rd-class-medical-be-eliminated/

Aircraft Owners and Pilots Association & Experimental Aviation Association (2012). Petition for exemption ...
FAA-issued medical certificate. Available from from http://www.regulations.gov/

Eidson, M. (2014). Example letter to Congressman or Senator. Retrieved from
http://www.supercub.org/forum/showthread.php?46526-Want-to-Know-What-the-AME-
Organization-Is-Saying-about-the-3rd-Class-Exemption

Federal Aviation Administration. (2008). *FAR Final Rule*. Retrieved from
https://www.leftseat.com/nprm_faa_medical_certificate_duration.htm

Goyer, R. (2014). AMA Vote Reveals Dirty Secret of 3rd Class Medical. *Flying Magazine*. Retrieved from
http://www.flyingmag.com/blogs/going-direct/ama-vote-reveals-dirty-secret-3rd-class-medical

Pilot's Bill of Rights 2, S.571. 114th Cong. (2015) Retrieved from https://www.congress.gov

Spence, C. (2014). Third Class Medical Reform Caught in Government Maze. *General Aviation News*. Retrieved
from http://generalaviationnews.com/2014/09/17/ third-class-medical-reform-caught-in-
government-maze/

Webb, J. (2014). Third Class Medical Letter to FAA Administrator. Retrieved from
http://www.asma.org/asma/media/asma/pdf-policy/2014/Third-Class-Medical-Letter-to-FAA-
Administrator-4-9-2014.pdf