Scientific Article

Blinding of Residency Applications in Medical Physics: Promises and Pitfalls

Ashley J. Cetnar, PhD,* Dominic J. DiCostanzo, PhD, Michael Weldon, MS, Amanda E. Swanson, MS, Joshua Mathews, PhD, and Nilendu Gupta, PhD

James Cancer Hospital and Solove Research Institute, The Ohio State University, Columbus, Ohio

Purpose: The feasibility of blinding applications for a medical physics residency program has yet to be demonstrated in the literature. We explore the application of an automated approach with human review and intervention to blind applications during the annual medical physics residency review cycle.

Methods and Materials: Applications were blinded using an automated process and used for the first phase of residency review in the program. We retrospectively compared self-reported demographic and gender data with blinded and nonblinded cohorts from 2 sequential years of review from a medical physics residency program. Demographic data were analyzed comparing applicants with candidates selected to move to the next phase of the review process. Interrater agreement was also evaluated from the applicant reviewers.

Results: We show the feasibility of blinding applications for a medical physics residency program. We observed no more than a 3% difference between the gender selection within the first phase of application review but greater differences when examining race and ethnicity between the 2 methods. The greatest difference was shown to be between Asian and White candidates, where there are statistical differences in the scores in the rubric categories of essay and overall impression.

Conclusions: We suggest that each training program critically evaluate its selection criteria for potential sources of bias within the review process. We recommend further critical investigation of processes to promote equity and inclusion to ensure the methods and outcomes are aligned with the mission of the program. Finally, we recommend that the common application provide an option for blinding applications at the source so this can be an option to facilitate efforts for evaluating unconscious bias in the review process.

Introduction

Program directors and admissions committees seek to minimize bias in assessing candidates in educational programs, but unconscious or implicit bias can affect hiring and recruiting decisions. Unconscious or implicit bias has been defined as "associations outside conscious awareness that lead to a negative evaluation of a person based on irrelevant characteristics such as race or gender."[1] Training programs have been developed to help reviewers become aware of these biases with varying levels of success.[2-6] In addition to education, a practical suggestion for

screening applicants is to "know as little about the candidate as possible" so reviewers do not form preconceived notions before assessing the candidate for the position because demographic information can potentially bias thoughts about an individual.[7,8]

Blinding applications is one strategy to help reduce unconscious bias in the application review process that removes or limits identifying information from the application to be used in the evaluation of a candidate.[9-11] Providing information about gender and ethnic origin may not be relevant for the job description or correlate with the achievements or potential of the candidate. Moreover, it could invoke unconscious biases during the review process.[7] Information that could be blinded in applications includes a candidate's name, photograph, educational background, and language.[7,11]

Several options are presented in the literature with methods to blind the recruitment processes. One method provides the option of viewing potentially biasing information in the review process by prompting the user with 2 questions: (1) whether they would like to view the data, such as name and photograph; and (2) whether they *should* view the data to help provide a checkpoint for reviewers.[9] Another option is to blind applications during the initial scoring of candidates and then provide unblinded information, with the option for score revision afterward.[9] However, when surveying human resources managers, less than 20% have reported using blinded hiring processes at their workplace.[9]

An example of success in blinding for gender diversity is the implementation of blind auditions for symphony orchestras, which effectively increased the probability of hiring and advancing female musicians within orchestras by concealing the musician's identity during the audition.[12] There is evidence to support that blinded interviews are an effective tool for improving interviews.[13] Fath et al have shown that "if job applications are stripped of identifying information, members of underrepresented social groups (ethnic minorities and women) become more likely to advance to the interview stage and, in certain cases, ultimately receive job offers."[9] Blinding applications presents a potential method for improving diversity and equity initiatives within teams or departments.

Within radiation oncology, physicians and medical physicists do not have equal representation when reviewing ethnicity and gender. Representation of Black and Hispanic doctors in radiation oncology is not proportional to the current proportions within the United States population at large.[14-16] If the 2021 climate survey responses serve as a representation of all American Association of Physicists in Medicine members, 2% of survey respondents were Black or African American (13.6% national representation) and 2% were Hispanic or Latino (18.9% nationally representation).[17] Limitations to this extrapolation include a 25% survey response rate, and 7%

of survey respondents did not respond to this demographic question. This is also true for women in radiation oncology and medical physics, where women comprise approximately 23% of medical physicists, whereas women represent 50.5% of the United States population.[18-21] In a study of a medical fellowship program, a cohort of interviewers were blinded to the written application and concluded there was no difference in the average rank overall as assigned by blinded compared with unblinded interviewers, but "blinded interviewers were more likely to rank underrepresented minority applicants higher."[22] In medicine, it has been reported that only 8% to 20% of residency programs use blinding during the recruitment process, but there are no known studies within medical physics assessing blinding during the application process.[23,24] Rationale and methods for blinding for medical physics residency applications has been introduced and implemented by other institutions.[25]

We explore the feasibility of implementing an automated approach with human review and intervention to blind applications during the annual medical physics residency review cycle. First, we retrospectively reviewed the results of blind candidate screening in the initial phase of our medical physics residency applicant review process. Then we compared the results to the previous year without blinding to assess how unconscious bias about the candidates could affect the screening methods. In addition, self-reported demographic data for gender and ethnicity were assessed by comparing candidate progress in the application process, comparing years where applications were and were not blinded. This study has been reviewed by The Ohio State University as an exempt study (#2022E0206).

Methods and Materials

Application review process

Each year, candidates apply for The Ohio State University Department of Radiation Oncology Medical Physics Residency Program through the Medical Physics Residency Application Program (MP-RAP) administered through the American Association of Physicists in Medicine. The recruitment of medical physics residents is the function of the Admissions Subcommittee of the Medical Physics Residency Education Committee (MPREC) within the department. The Medical Physics Admissions Subcommittee reviews the applications and related material in a 3-step process, which has been improved over time using an agile project management approach to make continued improvements each cycle.

Our institution currently uses a 3-phase approach to application review. In phase I of the review process, the applicant pool was ranked in the top half and bottom half

by teams of the Admissions Subcommittee members based on a full review of each application using a rubric for scoring. The top half of the applicants are invited for a short virtual interview in phase II. The Admissions Subcommittee meets to review the scores of all committee members and determine a score cut-off to use to invite candidates for a full virtual interview. The top-ranked candidates were invited to a final formal interview (phase III). Candidates are invited to a formal virtual interview by the members of the MPREC, with additional videos of the department provided as an orientation. To finalize decisions for those applicants invited to formal interviews, references are contacted by phone to supplement the recommendation letters, if needed.

In the unblinded year, 9 members of the Admissions Subcommittee, including physicists and senior medical physics residents, performed the phase I review. Within phase I, applicants were divided into 9 groups for review. Three reviewers were assigned for each of the applicant groups for full review using a standardized rubric for scoring topics of didactics, clinical exposure, references, essays, and overall impression. Approximately 10 to 12 applications were assigned per group. Reviewers were assigned to 3 separate applicant groups such that reviewers were paired into 3 different teams of reviewers. Within each group, approximately half of the applicants were recommended for phase II based on scores and consensus of reviewers.

In the blinded year, there were 10 members of the Admissions Subcommittee, including physicists and all medical physics residents. Six reviewers remained the same between years. With the introduction of blinding, the blinded applications were reviewed in the first phase, but unblinded information was provided before phase II.

Interviews are typically conducted in January and February. After completing all interviews, the Admissions Subcommittee met to review the interview results and other pertinent information regarding the applicants. The committee then internally ranks interviewees to determine the rank list for the MedPhys Match.

The Ohio State University participates in MedPhys Match, the medical physics residency-matching program for graduate students and postgraduate trainees. The internally ranked candidates are entered into the MedPhys Match by the published deadline for submitting the rank list. Upon completion of the Match process, offer letters are sent to successful candidates, as MedPhys Match requires. The interview and offer process is performed per the equal opportunity standards of The Ohio State University and Arthur G. James Cancer Hospital and Richard J. Solove Research Institute.

Blinding process

It is difficult to blind the provided applications manually or automatically based on the application information provided directly by MP-RAP. Our institutional process for blinding has included automating the download and renaming application PDF files. A Python script that uses the Selenium package and the Chrome web driver was developed to crawl the MP-RAP website and download each application. Selenium is a Python package specifically built to support the automation of web browsers. Developers are able to leverage Selenium to perform web scraping and automate rote tasks such as downloading multiple files in a table or capturing prices on e-commerce websites. After the PDFs were downloaded, another script was written to alphabetically sort each application by the applicant's last name and add a number (eg, 001, 002, 003) to the beginning of the file name. This method creates a file name structure (eg, "001_Doe, John.pdf") that supports parsing the data in each application. Data such as expected residency start date, highest degree, graduate schools, and American Board of Radiology certification status, among others, can be extracted from the structured data fields of the application and used to create a flat file database for the applicants in a comma-separated value format.

To anonymize the files, optical character recognition was performed in Adobe Acrobat for each application. This process takes approximately an hour to recognize the text in 100 applications. Once completed, a Python library was used for parsing PDFs.[26] A library of gendered pronouns and other words that would be static for all applications was generated, including male, female, he, she, him, her, hers, and his. Each word was surrounded by spaces (eg, " he "), which was critical to avoid erroneously blanking out partial words like the letters "he" in the word "the." The structured data fields of the application on the first pages containing identifying information were blinded for all applications through common identification of page regions. Since those pages are structured, the regions containing identifying applicant information were constant from one application to another, allowing for simple redaction of specific areas of the page.

The first step of the automated blinding process identified the applicant's name on the first page and stored the name to search for the name on all pages of the application. The search looped through every page to find the applicant's name and any of the gendered words specified. When these words were identified, the library applied a redaction by covering identified words with opaque rectangle objects. After redacting the application, the file was saved with a different file name, keeping only the application number to uniquely identify blinded applications and retain originals (eg, "001.pdf"). After the automated process was completed, the applications were reviewed by a member of the MPREC Admissions Subcommittee to identify gross errors introduced by the automation and identify applications requiring manual blinding. An example of the common areas of redaction are shown in Fig. 1.

Figure 1   An example of common areas of redaction within the resident applications.

Retrospective analysis

To demonstrate the feasibility of implementing this method, we retrospectively reviewed the demographic data from applicants to the medical physics residency program at The Ohio State University and applicant status through the applicant review process for 2 consecutive years: 1 unblinded, 1 blinded. The residency program has had approximately 100 applicants each year. Since this study focuses on underrepresented minority candidates, there is potential for applicants to be identified because of the small numbers within our field. To address the privacy of applicants in our study, the specific application years are not disclosed. The sample size is approximately 100 applicants each year, but exact numbers are not provided to protect the identity of the applicants. Percentages are used to present data for each year in aggregate.

Specific data analyzed includes self-reported gender (Female, Male, or prefer not to answer) and ethnicity from the common application, progress through review phases, and scores from reviewers. Within the MedPhys Match application system, applicants can provide primary ethnicity data and more specific information about their region of origin. For this study, applications were combined by the primary demographic identifier for 5 major categories: Asian, Black, Hispanic, White, and Prefer Not to Answer. Descriptive statistics were used to report gender and race and ethnicity data between cohorts. Applicants were assessed using a 7-point scale and rubric for each category, with at least 3 reviewers for each application. The difference between Asian and White applicant category scores was further investigated with a Student t test (2-tail).

The consistency of scores between reviewers, or interrater agreement, was evaluated for 4 evaluation categories: clinical experience, references, essays, and overall impression. Brown and Hauenstein's alpha coefficient[27] was chosen to determine the interrater agreement because it is independent of sample size and the distributions of scores within the rating scale. Alpha values range from [−1, 1], where values greater than or equal to 0.6 indicate moderate agreement, and values greater than or equal to 0.8 indicate strong agreement between reviewers. The correlations between scores in the individual categories (clinical experience, references, and essays) and the overall impression scores were evaluated using the Spearman correlation coefficient (r_s), which is a measure of directional covariance for variables in discrete ordinal data sets. Correlations for these categories were determined using the scores in each year (nonblinded and blinded) to determine whether any evaluation categories were strongly associated with reviewers' overall impressions of the applicants. The didactics category was excluded from these analyses because the academic performance criteria and scores did not vary between reviewers. Values of $r_s \geq 0.7$ indicate a significant correlation.

Results

The demographic data reported by the applicants are reported in Fig. 2, including the percentage of applicants and those passing through the phase I review where one year was not blinded and the next year was blinded. We observe similar selection rates when the process was unblinded, but differences specifically in the percentages of Asian and White applicants when blinding was implemented the next year. The selection rates for White applicants were 70% and 65% in years 1 and 2, respectively, and 56% and 41% for Asian applicants in years 1 and 2, respectively.

Reported genders of the applicants for the program are shown in Fig. 3, including the percentage of applicants
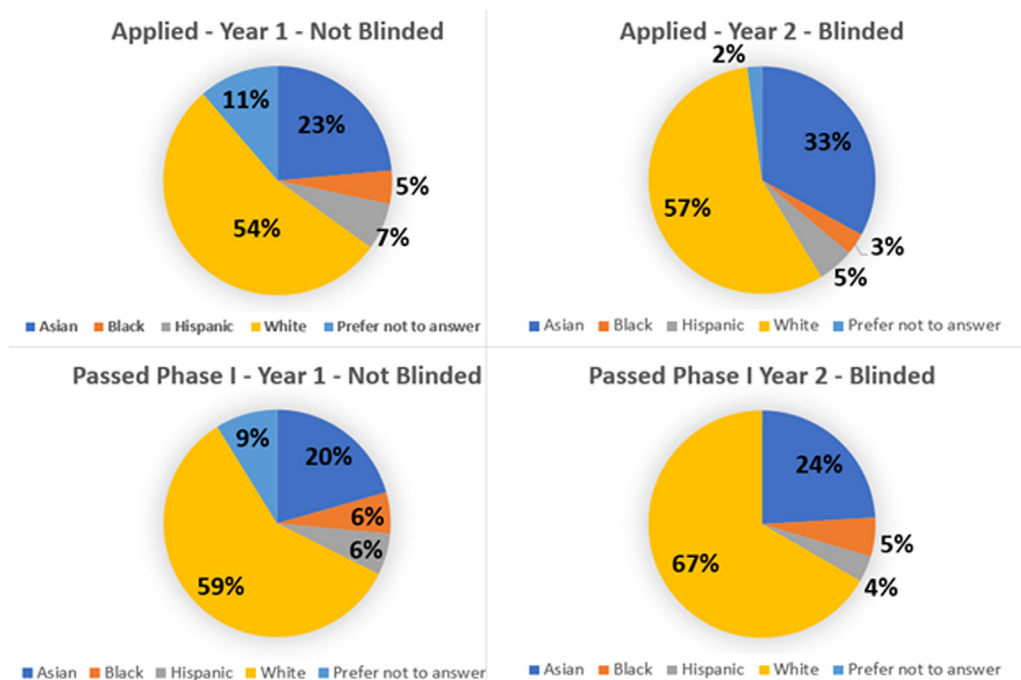
Figure 2 Comparison of race and ethnicity demographics for nonblinded (year 1) and blinded (year 2) applicant cohorts with approximately 100 applicants each year.

and those passing through the phase I review, where one year was not blinded and the next year was blinded, showing similar trends between methods that there is no preferential bias due to gender. Selection rates for male applicants were 64% and 55% in years 1 and 2, respectively, and for female applicants were 67% and 63% in years 1 and 2, respectively.

The changes in the overall percentage of each demographic category from total applicants through the phase I screening are shown in Table 1, showing differences primarily in Asian and White applicants between the 2 methods. The changes in the overall percentage of each gender category from total applicants through the phase I screening are shown in Table 1, showing no more than a
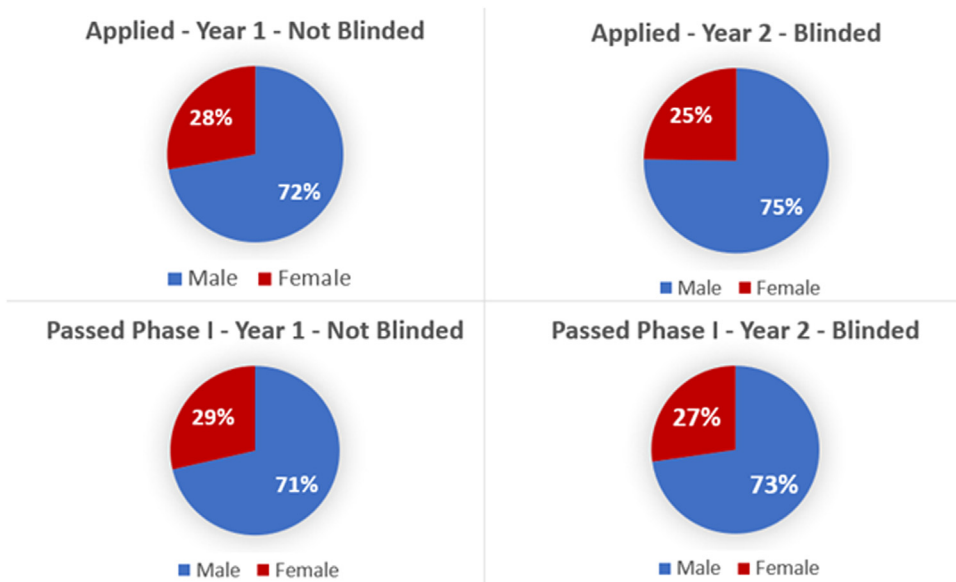

Figure 3 Comparison of gender between nonblinded (year 1) and blinded (year 2) applicant cohorts.

Table 1　Differences in gender and race and ethnicity between applicants admitted through phase I and total applicants for nonblinded (year 1) and blinded (year 2) cohorts

	Gender		Race and Ethnicity				
Year	Male	Female	Asian	Black	Hispanic	White	Prefer not to answer
1	−1%	1%	−3%	1%	−1%	5%	−2%
2	−3%	3%	−9%	2%	−1%	10%	−2%

Table 2　Average scores for Asian and White applicants during phase I for applications with statistical analysis using Student t test with bolded values denoted as statistically significant

	Didactics	Clinical exposure	References	Essays	Overall impression
Unblinded applications: Year 1					
Asian	4.60	3.97	4.00	3.55	3.72
White	4.75	4.53	4.15	3.93	4.06
P value	.698	.050	.438	**.025**	.147
Blinded applications: Year 2					
Asian	4.41	4.34	4.27	3.72	4.05
White	4.78	4.85	4.65	4.18	4.59
P value	.363	.057	.067	**.024**	**.009**

3% difference between the gender selection in phase I with either of the methods.

The difference between Asian and White applicants was further investigated to review the phase I scores shown in Table 2. We observe statistically significant differences in the scores in the essays and overall impression between Asian and White candidates.

Table 3 shows the percentage of applications demonstrating at least moderate or strong agreement between all reviewers for the blinded and nonblinded applicant pools. These data highlight that the references were scored most consistently among the categories (>95% of scores had at least moderate agreement between years), and the clinical exposure was scored least consistently (>75% of scores had moderate agreement). In addition, the interrater agreement for clinical exposure decreased in the blinded year (the percentage of scores with at least moderate agreement decreased from 81% to 75%), whereas the agreement for the remaining categories was relatively consistent between the blinded and nonblinded years.

Figure 4 shows the correlation coefficients for the blinded and nonblinded years. Using all reviewer scores, the references were most strongly correlated to the overall impression, and the clinical exposure scores were least correlated to the overall impression for both years. Correlations varied for different reviewers, but only the reference category significantly correlated with the overall impression using all scores in the blinded year.

Discussion

The goal of blinding was to reduce unconscious bias in the application review process by removing identifying

Table 3　Percentage of application scores that showed various levels of interrater agreement between all reviewers for the unblinded and blinded cohorts*

	Year 1: Not blinded				Year 2: Blinded			
Category	Clinical exposure	References	Essay	Overall	Clinical exposure	References	Essay	Overall
Strong agreement	51%	81%	71%	76%	48%	71%	66%	65%
Moderate agreement	81%	95%	90%	94%	76%	96%	91%	90%
Remainder (poor agreement)	19%	5%	10%	6%	24%	4%	9%	10%

* Interrater agreement was evaluated using Brown and Hauenstein's alpha coefficient.
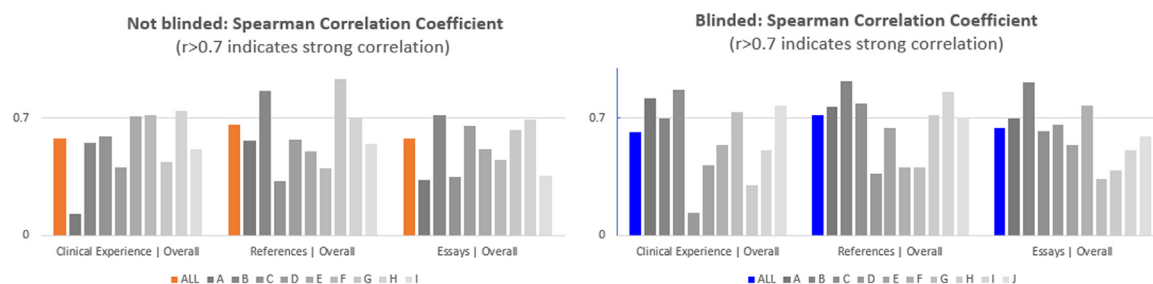
Figure 4 Spearman correlation coefficients for clinical experience, references, and essays compared with the overall impression for the nonblinded and blinded applicant pools. A horizontal bar for $r_s = 0.7$ is included in both plots to demarcate the threshold for significant correlation. Reviewers are labeled by letter, and label order is not preserved between years.

information from candidate applications. By introducing the blinding method to our process, we preserved the selection rates between initial applications and those passing through the blinded phase 1 review for gender. Although our institution has shown similar selection rates between genders in this study, other studies have shown gendered names as a significant variable for evaluating candidates.[28-31] Although there is no parity in gender for the applicants during these review cycles, validating this selection criterion supports the evaluation equity of candidate qualification between genders.

Although we do not observe large differences in the gender composition between screening methods, we observe differences in the race and ethnicity between applicants and those admitted through phase I of the screening process. As shown in Table 1, we see approximately a 10% decrease in Asian applicants and an increase of approximately 10% in White applicants between the total applicants and those admitted through the phase I screening. In addition, there are statistically significant differences in the scores in the essays and overall impression between Asian and White candidates in our blinding cohort. Some hypotheses for the differences in scores include cultural differences in projecting accomplishments and credentials in curricula vitae, possible unequal representation of nonnative English writers for personal statement submissions, and no explicit scoring of research experience in the phase I review for the clinically focused residency program. This study currently does not investigate reasons for these differences, but additional data, such as understanding backgrounds and pathways of individual applicants instead of considering demographic information in aggregate cohorts, would be of interest for future work.

Another interesting observation from the data is that many applicants choose not to identify their demographic information, demonstrating the desire to self-blind gender and other demographic details from the institution. Within our cohorts, the percentage of applicants choosing to self-blind decreased from 11% to 2% between the 2 years. However, this accounts for a significant

percentage of applicants not considered in this study's secondary analysis. This trend has been shown in other medical physics surveys with reasons suggested, including a lack of trust in how the data will be used or insufficient options available to be able to represent themselves.[17]

In our work, we investigated the feasibility of blinding the first of 3 phases of our institutional process, but blinding could be adapted to others recruiting for their residency programs. Blinding applications presents a potential method for improving diversity and equity initiatives by considering candidates based on qualifications. Unconscious bias could negatively affect the screening of applications, which is the initial step impacting the representation of diverse candidates who are interviewed for a position and ultimately selected within programs. There is concern that even if applications are blinded during the initial review, discrimination may be postponed to later in the process.[7] It has been shown that underrepresented minority applicants may experience bias in aspects such as letters of recommendation, fewer research opportunities, and lower test scores.[7,22] We recommend that programs use a standardized rubric for assessing applications to ensure consistency in evaluation, crucially assessing both the metrics and weighting within the rubric, since scoring could disproportionately affect minority candidates. Additional techniques beyond traditional application review and interviews could be expanded to allow candidates to demonstrate skills or participate in sample projects to allow an opportunity for underrepresented minorities to be competitive.[11]

Although there is promise to blinding applications to minimize unconscious bias, the challenges must also be discussed when applying this methodology. Implementing blinding for medical physics applicants through the MP-RAP system is currently not an option for programs, so there is a burden of time and scripting to develop in-house methods for blinding the applications. As a result, this process is resource intensive, and not all facilities can straightforwardly implement these methods if desired. We recommend that the common application provide an option for blinding applications at the source so this can

be an option to facilitate efforts for evaluating unconscious bias in the review process. In addition, changes in application format are minimized from year to year, and changes to the application are announced with examples provided before the opening of the application to allow for prospective adaptation of software to new application formats.

The automated process generated challenges and was not seamless. It required additional human-directed post-processing and quality control. When identifying other gendered pronouns, instead of using "he," " he " with spaces surrounding the word was used during the automated process. There was a problem identifying "Ms" as a gendered title versus the educational qualification of a master of science (MS) within the applications. "Ms" was used as the specific search criteria to ensure it was only the salutation. Names that are similar to words or are very short (2-3 letters) can result in the word being eliminated (eg, "Li" resulted in "linac" becoming "nac"). Being specific in searches for spaces and capitalization can reduce this. If names/pronouns are removed entirely, it can be challenging to read reference letters and personal statements. This can be adapted by replacing pronouns with neutral equivalent versions. Nicknames, middle names, or other versions of candidates' names that are different than their names stated in the application may not be redacted if they are used. These have to be identified and addressed manually.

Some individual applications were no longer comprehensible as a result of the redactions because of the watermarks on transcripts or background images included on letters of recommendation. PDFs within the application with layers/watermarks may cause issues with automated redaction methods, where entire pages are removed. Photos that are included in documents (mainly resumes/curricula vitae) that are uploaded are still present with current methods. These are examples of limitations that still need to be addressed manually. Previous employment or names of schools outside of the United States may result in reviewers making assumptions about the candidate despite removal of names/pronouns. The explicit disclosure of languages spoken by the candidate could also provide information that could bias reviewers.

Limitations to this study include the retrospective study from a single institution with 1 year for blinded and nonblinded comparison. However, it demonstrates the feasibility of analyzing data that could be replicated for other programs to implement in evaluating their process of analyzing unconscious bias. By critically evaluating the process for screening applicants, we can continue to improve institutional processes. In our phase I scoring rubric, we reviewed the interrater agreement for our admissions committee and the value of our current scoring categories within the rubric. Since the correlation between the clinical, reference, and essay categories and

the overall impression increased in the blinded year, perhaps the overall impression was influenced by demographic factors in the nonblinded cohort that were not available in the blinded cohort. After the automated review, an Admissions Subcommittee member performed the final screening of the applications. This reviewer could be influenced by the prior knowledge of the applications but was one of several reviewers for each application to minimize this potential bias. A limitation to self-reporting is that the responses may not be accurate. Although there is a choice for participants not to disclose information, an assumption for this study is that the reported gender and ethnicity of participants is a correct representation of the applicants and this data was not verified by another method.

In summary, the potential advantages of blinding resident applications include the following: (1) reduced bias based on gender, race, and ethnicity; and (2) maintaining diversity of applicants throughout phases of review. Potential disadvantages include the following: (1) challenges in remembering candidates after evaluating them when names are removed; (2) in some cases, blinding can make essays hard to read if names and pronouns are not replaced with neutral versions; (3) automated processes are not perfect and may result in extracting words similar to names that were extracted; and (4) editing is time consuming.

Conclusion

In this study, we show the feasibility of blinding applications for a medical physics residency program. We suggest that each training program critically evaluate its selection criteria for potential sources of bias within the review process. We recommend further critical investigation of processes to promote equity and inclusion to ensure the methods and outcomes are aligned with the mission of the program. Finally, we recommend that the common application provide an option for blinding applications at the source so this can be an option to facilitate efforts for evaluating unconscious bias in the review process.

Acknowledgments

References

1. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: A systematic review. *BMC Med Ethics*. 2017;18:19.

2. Jackson SM, Hillard AL, Schneider TR. Using implicit bias training to improve attitudes toward women in STEM. *Soc Psychol Educ*. 2014;17:419-438.

3. Jackson JL. The non-performativity of implicit bias training. *Radical Teacher*. 2018;112:46-54.

4. Applebaum B. Remediating campus climate: Implicit bias training is not enough. *Stud Philos Educ*. 2019;38:129-141.

5. Pritlove C, Juando-Prats C, Ala-Leppilampi K, Parsons JA. The good, the bad, and the ugly of implicit bias. *Lancet*. 2019;393:502-504.

6. Sherman M, Ricco J, Nelson S, Nezhad S, Prasad S. Implicit bias training in a residency program: Aiming for enduring effects. *Fam Med*. 2019;51:677-681.

7. Kumar RK. Blind hiring: A solution to BIAS. *IIBM J Manag Res*. 2018;3:83-89.

8. Huffcutt AI. From science to practice: Seven principles for conducting employment interviews. *Appl HRM Res*. 2010;12:121-136.

9. Fath S, Larrick RP, Soll JB, Zhu S. Why putting on blinders can help us see more clearly. *MIT Sloan Manag Rev*. 2021;62:38-45.

10. Miles WS, Shaw V, Risucci D. The role of blinded interviews in the assessment of surgical residency candidates. *Am J Surg*. 2001;182:143-146.

11. Maurer R. Blind hiring may be missing the point. Available at: https://www.shrm.org/resourcesandtools/hr-topics/talent-acquisition/pages/blind-hiring-practices.aspx. Accessed January 11, 2023.

12. Goldin C, Rouse C. Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*. 2000;90:715-741.

13. McDaniel MA, Whetzel DL, Schmidt FL, Maurer SD. The validity of employment interviews: A comprehensive review and meta-analysis. *J Appl Psychol*. 1994;79:599-616.

14. Lightfoote JB, Fielding JR, Deville C, et al. Improving diversity, inclusion, and representation in radiology and radiation oncology part 1: Why these matter. *J Am Coll Radiol*. 2014;11:673-680.

15. Nead KT, Linos E, Vapiwala N. Increasing diversity in radiation oncology: A call to action. *Adv Radiat Oncol*. 2019;4:226-228.

16. Lightfoote JB, Fielding JR, Deville C, et al. Improving diversity, inclusion, and representation in radiology and radiation oncology part 2: Challenges and recommendations. *J Am Coll Radiol*. 2014;11:764-770.

17. Hendrickson KR, Avery SM, Castillo R, et al. 2021 AAPM equity, diversity, and inclusion climate survey executive summary [e-pub ahead of print]. *Int J Radiat Oncol Biol Phys*. 2021. https://doi.org/10.1016/j.ijrobp.2022.02.030, accessed January 11, 2023.

18. Chapman CH, Hwang W-T, Deville C. Diversity based on race, ethnicity, and sex, of the US radiation oncology physician workforce. *Int J Radiat Oncol Biol Phys*. 2013;85:912-918.

19. Covington EL, Moran JM, Paradis KC. The state of gender diversity in medical physics. *Med Phys*. 2020;47:2038-2043.

20. van Zyl M, Haynes EM, Batchelar D, Jakobi JM. Examining gender diversity growth as a model for inclusion of all underrepresented persons in medical physics. *Med Phys*. 2020;47:5976-5985.

21. United States Census Bureau. Quick facts. Available at: https://www.census.gov/quickfacts/fact/table/US/LFE046221. Accessed January 11, 2023.

22. Haag J, Sanders B, Keach JW, Lefkowits C, Sheeder J, Behbakht K. Impact of blinding interviewers to written applications on ranking of Gynecologic Oncology fellowship applicants from groups underrepresented in medicine. *Gynecolog Oncol Rep*. 2022;39:100935.

23. Kasales C, Peterson C, Gagnon E. Interview techniques utilized in radiology resident selection—A survey of the APDR. *Acad Radiol*. 2019;26:989-998.

24. Kim RH, Gilbert T, Suh S, Miller JK, Eggerstedt JM. General surgery residency interviews: Are we following best practices? *Am J Surg*. 2016;211:476-481.

25. Paradis KC, Covington EL. Breaking down roadblocks to diversity: Managing bias in hiring. *Am Assoc Phys Med Newsletter*. 2021;46:49-50.

26. GitHub. PyMuPDF. Available at: https://github.com/pymupdf/PyMuPDF. Accessed January 11, 2023.

27. Brown RD, Hauenstein NM. Interrater agreement reconsidered: An alternative to the rwg indices. *Organ Res Methods*. 2005;8:165-184.

28. Moss-Racusin CA, Dovidio JF, Brescoll VL, Graham MJ, Handelsman J. Science faculty's subtle gender biases favor male students. *Proc Natl Acad Sci*. 2012;109:16474-16479.

29. Steinpreis RE, Anders KA, Ritzke D. The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles*. 1999;41:509-528.

30. Reuben E, Sapienza P, Zingales L. How stereotypes impair women's careers in science. *Proc Natl Acad Sci*. 2014;111:4403-4408.

31. Quadlin N. The mark of a woman's record: Gender and academic performance in hiring. *Am Sociol Rev*. 2018;83:331-360.