

# Misalignment of Biostatistics Content Between Licensing Exam Study Aids and Contemporary Medical Research

W. Connor Haycox, MD<sup>a</sup>; Dmitry Tumin, PhD<sup>b</sup>

## AUTHOR AFFILIATIONS:

<sup>a</sup>Department of Family Medicine, Brody School of Medicine, East Carolina University, Greenville, NC

<sup>b</sup>Department of Academic Affairs, Brody School of Medicine, East Carolina University, Greenville, NC

## CORRESPONDING AUTHOR:

W. Connor Haycox, Department of Family Medicine, Brody School of Medicine, East Carolina University, Greenville, NC, [haycoxc24@ecu.edu](mailto:haycoxc24@ecu.edu)

**HOW TO CITE:** Haycox WC, Tumin D. Misalignment of Biostatistics Content Between Licensing Exam Study Aids and Contemporary Medical Research. *Fam Med*. 2024;57(X):1–6. doi: [10.22454/FamMed.2024.967125](https://doi.org/10.22454/FamMed.2024.967125)

**PUBLISHED:** 5 December 2024

**KEYWORDS:** statistics, teaching about research, USMLE

© Society of Teachers of Family Medicine

## ABSTRACT

**Background and Objectives:** Medical trainees express difficulty with interpreting statistics in clinical literature. To elucidate educational gaps, we compared statistical methodologies in biomedical literature with biostatistical content in licensing exam study materials.

**Methods:** In this bibliographic content analysis, we compiled a stratified random sample of articles involving original data analysis published during 2023 in 72 issues of three major medical journals. We recorded all discrete statistical methods and concepts detailed in the methods section of the articles and in three commercial licensing exam study resources. We created a unified list of discrete methods or concepts to define overarching domains and mapped each method to a domain to determine that domain's presence in each resource or article.

**Results:** In a sample of 273 journal articles and three study resources, we identified 1,057 unique key words mapped onto 20 domains. Statistical error, significance, power analysis, and group comparisons of categorical data were high-frequency domains among the articles. Overall, 63% of articles included methods from domains not covered in any study resource.

**Conclusions:** Medical licensing exam preparation does not reflect the breadth of contemporary statistics in biomedical research. Future interventions should expand medical students' understanding of research protocols and complex data manipulation.

## INTRODUCTION

Statistical literacy is a key skill set developed during medical training that facilitates physicians' evidence-based clinical reasoning and application of scientific literature.<sup>1–3</sup> Physicians' application of biostatistics knowledge supports their ability to interact knowledgably with the literature (including both primary research and evidence-based guidelines) and to make evidence-based decisions regarding the care of their patients. Primary care represents many patients' sole contact with medical professionals, so for family physicians to be able to effectively interpret clinical research to help their patients make informed decisions regarding their health is especially important.

Medical students now commonly complete undergraduate biostatistics coursework to develop these skills prior to medical school matriculation,<sup>4</sup> and preclinical curricula now increasingly emphasize exposure to concepts of evidence-based medicine concepts.<sup>5–10</sup> In addition to the clinical relevance of comprehensive biostatistics training, biostatistics and epidemiology are core components of national licensing

examinations for medical trainees.<sup>11,12</sup> Given the importance of licensing exam performance for residency placement, many medical students supplement formal curricular content with self-guided commercial resources, including texts, videos, question banks, and preparatory courses.<sup>13,14</sup>

Despite exposure to biostatistical content in the undergraduate medical education curriculum and in cocurricular commercial study resources, current medical students and practicing physicians continue to express difficulty with interpreting statistics in the clinical literature, especially in studies that use complex or novel statistical methods.<sup>15–17</sup> This challenge has been exacerbated by the increased complexity of statistical methods used in the medical literature, creating a mismatch between advancing biostatistical methods in medical research and the relatively static content of study resources and curricula,<sup>18–21</sup> and prompting calls for enhanced biostatistics training in medical education.<sup>18,22</sup> The ubiquity of commercial study resources for licensing exams means that medical students frequently are exposed to biostatistics content through this medium, so understanding whether bio-

statistics knowledge acquired from exam study aids would plausibly prepare students to interpret biostatistical methods in contemporary medical research is important. Moreover, analysis of the biostatistics content in these resources could reveal potential deficits in statistics education and inform future redesigns of curricula and study materials.

In this study, we aimed to compare statistical methods and concepts employed in contemporary biomedical literature to biostatistical topics covered in commercial study resources used to prepare for licensing examinations taken by medical students in the United States—specifically, United States Medical Licensing Examination (USMLE) Step 1 and COMLEX–USA Level 1. We hypothesized that a significant discrepancy exists between the biostatistical methods described in board study materials and those used in contemporary medical literature, including (a) an emphasis in licensing exam study materials on biostatistics concepts that are rarely encountered in the biomedical literature, and (b) a lack of coverage of other biostatistics concepts that are frequently encountered in contemporary biomedical research.

## METHODS

This was a bibliographic study not involving human research subjects and therefore did not require institutional review board approval. Following similar content analysis studies,<sup>18,22</sup> we sampled issues of the *Journal of the American Medical Association (JAMA)*, the *British Medical Journal (BMJ)*, and the *New England Journal of Medicine (NEJM)* using a stratified random sampling approach. Based on weekly issues published during 2023, we randomly selected two issues per month per journal, yielding a total of 72 issues. Within each issue, we selected articles for review if they involved primary or secondary data analysis, including original research (clinical or basic science), brief reports or research letters, and systematic reviews with meta-analyses. We excluded letters (excepting research letters reporting original data analysis), images, commentaries, case reports, editorials, treatment guidelines, and any other article type not reporting results from original data analysis. We also excluded theme issues; if a theme issue was selected randomly, we used the subsequent available issue.

We determined a reading schedule a priori by randomly sorting issue and article order, and then reviewed each article for statistical content. We recorded each discrete statistical method (eg, linear regression model) or concept (eg, *P* value) described in the methods section of the main article text. Although biostatistics often are taught to medical students alongside epidemiological and other concepts, we focused our study on statistical methods only, operationalized as methods used to summarize, compare, or interpret quantitative data. For each article, we recorded the presence of each statistical method or concept as a binary variable. To establish interrater agreement, two authors read and analyzed the first 20 articles. Thereafter, a single author reviewed all remaining articles and extracted key words or phrases that were used in the final analysis. Any coding discrepancies in the first 20 articles and any ambiguous coding of subsequent articles were resolved via

discussion and consensus.

We selected USMLE Step 1 and COMLEX–USA Level 1 commercial study aids based on medical students' documented use of these resources and positive correlation with licensing exam scores.<sup>3–6</sup> To incorporate a wide range of affordable materials corresponding to varied study modalities, we chose one comprehensive review text, one online biostatistics-specific question bank, and an online review video depository, all updated in 2023. Given the relatively succinct nature of these materials, one author reviewed the entirety of each resource's biostatistics content.

After completing data extraction, we created a unified list of all discrete methods or concepts (key words or phrases) extracted from the sampled articles and study resources. Two authors independently generated a list of domains encompassing these methods and concepts, and they arrived at a final domain list after discussion. The two authors then independently mapped each method or concept mentioned in the articles and study resources to one of the domains, and they arrived at a final classification after discussion of discrepant domain coding. For each domain, we determined (a) whether it was present or absent in each study resource; (b) whether it was present or absent in each journal article; and (c) in how many articles it was present out of the total number of sampled articles (frequency). A priori, we defined “frequently occurring” concepts in the biomedical literature as categories in the top quartile of frequency and “infrequently occurring” categories as those in the bottom quartile of frequency. Complete data and replication code for our analyses are included as Supplemental Files 1 and 2.

## RESULTS

We identified 273 eligible journal articles (124 in *JAMA*, 53 in *BMJ*, and 96 in *NEJM*) and extracted 4,728 key words from these articles and from the three study resources. After excluding 776 key words determined not to represent a specific statistical method (eg, excluding the phrase “sensitivity analysis,” which could refer to any method or test depending on the context) and removing duplicates in the remaining key words, we retained 1,057 unique key words (methods or concepts) for further classification into domains of biostatistical methods. We summarized the biostatistical methods and concepts into 19 domains and one residual category of statistical methods not elsewhere classified (Table 1).

We examined coverage of each domain within the three board exam study resources (Table 2). These study resources lacked any coverage of methods related to comparing non-normally distributed data between groups, fitting generalized linear regression or multilevel regression models, or using weights in statistical analysis. Additionally, the board exam study resources did not always cover concepts related to linear regression, survival analysis, meta-analysis, or handling of missing data. Importantly, Table 2 represents coverage of any concepts within each domain but does not address the depth of this coverage. For example, all study resources were

**TABLE 1.** Domains of Biostatistical Methods Found in Journal Articles or Study Resources

Domain	Examples of words or phrases mapped to domain
1. Summary statistics: Measures of central tendency	Mean, median
2. Summary statistics: Measures of dispersion	Standard deviation, interquartile range
3. Bivariate comparisons: Group comparison of categorical data	Relative risk, $\chi^2$ test
4. Bivariate comparisons: Group comparison of normally distributed continuous data	<i>t</i> test, mean difference
5. Bivariate comparisons: Group comparison of nonnormally distributed continuous data	Rank sum test, Kruskal-Wallis test
6. Bivariate comparisons: Correlation of continuous data	Correlation, Pearson correlation coefficient
7. Statistical significance and multiple testing	<i>P</i> value, $\alpha$
8. Statistical error, variance, and confidence intervals	Confidence interval, standard error
9. Power analysis, sample size calculation, and stopping criteria	Power, interim analysis
10. Survival analysis: Univariate and bivariate methods	Kaplan-Meier, log-rank test
11. Survival analysis: Multivariable regression methods	Hazard ratio, Cox proportional hazards model
12. Multivariable regression (excluding survival analysis): Linear regression	Linear regression, $R^2$
13. Multivariable regression (excluding survival analysis): Generalized linear regression	Logistic regression, Poisson regression
14. Multivariable regression (excluding survival analysis): Multilevel regression	Fixed effects, random effects
15. Multivariable regression (excluding survival analysis): Effect modification and mediation	Interaction term, effect modifier
16. Diagnostic test and predictive score performance	Sensitivity, specificity
17. Missing data and imputation of missing values	Imputation, missingness
18. Meta-analysis	Heterogeneity, forest plot
19. Weighting and complex sampling design	Weighting, inverse probability weighting
20. Statistical methods not elsewhere classified	Log transformation, cubic spline

credited with covering the “statistical significance and multiple testing” domain because all three resources addressed *P* values,  $\alpha$  levels, and hypothesis testing, but none of these resources went so far as to discuss correction of significance testing for multiple comparisons.

We compared the frequency of each domain within the 273 journal articles included in our analysis (Table 3). High-frequency domains included statistical error, variance, and confidence intervals; statistical significance and multiple testing; measures of central tendency; power analysis, sample size calculation, and stopping criteria; and group comparisons of categorical data. Low-frequency domains included correlation of continuous data; group comparison of nonnormally distributed continuous data; diagnostic test and predictive score performance; meta-analysis methods; and weighting. Considering the four domains that received no coverage in any of the three board exam study resources (as shown in Table 3), we found that 172 of the 273 (63%) articles included methods from at least one of these domains.

## DISCUSSION

The accelerating pace of discovery in biomedical research highlights the importance of interpreting original research in light of emerging diseases, new therapeutic options, and changes in population health.<sup>18,19</sup> Primary care physicians are uniquely positioned to interpret and apply the findings of biomedical research because they are often patients’ primary or even only point of contact within the medical system. Despite the availability of evidence-based point-of-care resources synthesizing the research literature for practical use, the many

instances of rapidly evolving science, conflicting evidence, or conflicting interpretations of the same evidence should prompt physicians to examine the methods and findings of the primary literature more closely. In this study, we demonstrated that a major discrepancy persists between biostatistical methods frequently occurring in clinical research and the methods that are emphasized in medical licensing exam study resources, potentially contributing to significant learning gaps among physicians in training. Our study thus reinforces prior studies’<sup>22</sup> calls for overhauling statistics curricula in undergraduate medical education and provides initial data regarding specific statistical methods to be incorporated in revised curricula.

One-fifth of all domains of biostatistical methods were omitted across all three study resources, including multilevel regression modeling and data weighting. Notably, all but one of the domains (ie, generalized linear regression) defined in our study were represented in the most recent USMLE and COMLEX-USA content outlines, which encompass expected knowledge across all three standardized licensing exams.<sup>11,12</sup> All high-frequency domains noted in the journal articles were included in all study resources, but their coverage was often superficial relative to the articles’ level of detail. The complexity of a statistical method (from the reader’s point of view) is based on the level of prior knowledge assumed when describing the method, the computational difficulty of implementing that particular method, and the relevance of that method to the reader’s area of practice, with different statistical methods being more or less common in different specialties. Our coding scheme assumed that if a student learns

**TABLE 2.** Representation of Domains of Biostatistical Methods in Study Resources for Board Exams

Domain	Study resource		
	A	B	C
1. Summary statistics: Measures of central tendency	X	X	X
2. Summary statistics: Measures of dispersion	X	X	X
3. Bivariate comparisons: Group comparison of categorical data	X	X	X
4. Bivariate comparisons: Group comparison of normally distributed continuous data	X	X	X
5. Bivariate comparisons: Group comparison of nonnormally distributed continuous data			
6. Bivariate comparisons: Correlation of continuous data	X	X	X
7. Statistical significance and multiple testing	X	X	X
8. Statistical error, variance, and confidence intervals	X	X	X
9. Power analysis, sample size calculation, and stopping criteria	X	X	X
10. Survival analysis: Univariate and bivariate methods	X		X
11. Survival analysis: Multivariable regression methods	X		X
12. Multivariable regression (excluding survival analysis): Linear regression	X		X
13. Multivariable regression (excluding survival analysis): Generalized linear regression			
14. Multivariable regression (excluding survival analysis): Multilevel regression			
15. Multivariable regression (excluding survival analysis): Effect modification and mediation	X	X	X
16. Diagnostic test and predictive score performance	X	X	X
17. Missing data and imputation of missing values	X		X
18. Meta-analysis	X		
19. Weighting and complex sampling design			
20. Statistical methods not elsewhere classified	X	X	X

**TABLE 3.** Representation of Domains of Biostatistical Methods in Journal Articles

Domain	Number of articles
1. Summary statistics: Measures of central tendency	153
2. Summary statistics: Measures of dispersion	107
3. Bivariate comparisons: Group comparison of categorical data	139
4. Bivariate comparisons: Group comparison of normally distributed continuous data	68
5. Bivariate comparisons: Group comparison of nonnormally distributed continuous data	25
6. Bivariate comparisons: Correlation of continuous data	30
7. Statistical significance and multiple testing	222
8. Statistical error, variance, and confidence intervals	227
9. Power analysis, sample size calculation, and stopping criteria	141
10. Survival analysis: Univariate and bivariate methods	88
11. Survival analysis: Multivariable regression methods	91
12. Multivariable regression (excluding survival analysis): Linear regression	50
13. Multivariable regression (excluding survival analysis): Generalized linear regression	115
14. Multivariable regression (excluding survival analysis): Multilevel regression	79
15. Multivariable regression (excluding survival analysis): Effect modification and mediation	57
16. Diagnostic test and predictive score performance	15
17. Missing data and imputation of missing values	101
18. Meta-analysis	40
19. Weighting and complex sampling design	36
20. Statistical methods not elsewhere classified	192

about statistical power, for example, they also understand interim analyses and stopping criteria derived from power analyses. However, incomplete coverage of these concepts in study resources can contribute to medical graduates' limited ability<sup>23–26</sup> and overestimation<sup>27,28</sup> of their skills to use statistics to make evidence-based decisions. While all study resources emphasized concepts relating to diagnostic testing and predictive score performance (eg, sensitivity, specificity), this major component of licensing exam preparation rarely was encountered in the journal articles, though we acknowledge its distinct use in clinical practice.<sup>29,30</sup> Nonetheless, even with perfect understanding and retention of biostatistics content encountered on licensing exams, medical trainees may not be prepared to interpret contemporary statistical methodology in biomedical research.<sup>31</sup>

Prior studies in medical education have indicated a need for improved statistical training,<sup>32</sup> and our study underscores the need for enhancing education in this area. Given the breadth of statistical concepts, expecting medical students to become familiar with every methodological term that might appear in primary research literature is unrealistic; rather, educators must find effective frameworks for maximizing statistical learning of core concepts and frequently used techniques during the preclinical and clinical years.<sup>33</sup> Within medical school curricula, enhanced longitudinal exposure with spaced repetition may be the first step, along with journal clubs and research review during case presentations.<sup>7</sup> Looking ahead to these interventions in the clinical years, the foundational curriculum also can signpost statistical concepts that likely will remain relevant to students' education and practice beyond the period of preparing for licensing exams. Faculty development to improve the teaching of statistics can be supplemented by hiring biostatistician educators, although this may be difficult at schools without affiliated biostatistics departments.<sup>9</sup> Study resource revision may be the most challenging area to reform directly due to the independence of third-party companies producing study resource content; but faculty- and student-led efforts to create study materials can be scaled up to supplement commercial resources.<sup>34</sup> Regardless of what changes medical educators ultimately make, intentionality and persistence are prerequisites for improving biostatistical education.

The present study is limited in several regards. Analysis of commercial study resources may not reflect the content of medical school curricula, which vary across institutions. Furthermore, the authors' professional background and experience with statistical analysis influenced the development of the coding scheme, and other groups may prefer more granular schemes or different classifications of specific terminology within the domains presented in our study. The three commercial study resources were chosen for their varied modalities and students' documented preferences, but they may not be fully representative of all available study aids for licensing exams. Similarly, we chose to sample articles from three high-impact journals selected for their wide readership and reputation, but our findings may not be generalizable to all recent biomedical

research, especially because we sampled articles from only 1 year. Moreover, we extracted data on statistical concepts from the methods section of each published article, but did not review the results, tables, figures, or supplemental content, which may have contained additional statistical terminology not mentioned in the methods section.

Overall, our comparison of biostatistics content in medical licensing exam study resources and three major biomedical journals indicates that medical trainees may be critically underprepared to interact with contemporary clinical research because of incomplete coverage of statistical methods in materials aimed at preparing them for standardized tests. Limited longitudinal training in biostatistics methods can restrict students' ability to critically evaluate both primary research and point-of-care resources that synthesize the existing literature, as well as potentially limit their ability to contribute to primary care research once they are in practice. In addition to testing the efficacy of educational interventions to close this gap, future studies could incorporate content analysis of medical school statistical curricula; expand the sample size and date range of sampled articles and study resources; examine how study resources are matched or mismatched with biostatistical methods that are most common in particular specialties; or conduct longitudinal analyses of multiple editions of study resources to assess for possible evolution of statistical content. We must ensure that medical students and physicians become better trained to recognize and interpret the wide array of statistical methodologies that legitimize or invalidate the novel diagnostic and treatment options offered to patients.

### Presentations

This article's content was presented in poster form at the Brody School of Medicine Distinction Day (Greenville, NC; April 11, 2024) and in a brief podium presentation at the Brody School of Medicine Medical Education Day (Greenville, NC; April 25, 2024).

### Conflict Disclosure

D.T. discloses salary support from Kate B. Reynolds Charitable Trust and Lilly Grant Office for research and quality improvement projects unrelated to this work. The authors have no other competing interests to declare.

### REFERENCES

1. Rao G, Kanter SL. Physician numeracy as the basis for an evidence-based medicine curriculum. *Acad Med*. 2010;85(11):799.
2. Oster RA, Enders FT. The importance of statistical competencies for medical research learners. *J Stat Educ*. 2018;26(2):137–142.
3. Schmidt FM, Zottmann JM, Sailer M, Fischer MR, Berndt M. Statistical literacy and scientific reasoning & argumentation in physicians. *GMS J Med Educ*. 2021;38(4):77.
4. Medical School Admission Requirements™ (MSAR®) Report for Applicants and Advisors. *Association of American Medical Colleges*. 2023. <https://students-residents.aamc.org/system/files/2023-07/>

[MSAR\\_Premed\\_Course\\_Requirements\\_07.17.23.pdf](#).

5. Arps K, Schulman D, Tigges S. Clinical research statistics: an interactive self-study quiz. *MedEdPORTAL*. 2014;10:9808.
6. Evans KH, Thompson AC, Brien O, C. An innovative blended preclinical curriculum in clinical epidemiology and biostatistics: impact on student satisfaction and performance. *Acad Med*. 2016;91(5):696–700.
7. Swanberg SM, Mi M, Engwall K. An integrated, case-based approach to teaching medical students how to locate the best available evidence for clinical care. *MedEdPORTAL*. 2017;13:10531.
8. O'neil J, Croniger C. Critical appraisal worksheets for integration into an existing small-group problem-based learning curriculum. *MedEdPORTAL*. 2018;14:10682–10682.
9. Gold JG, Knight CL, Christner JG, Mooney CE, Manthey DE, Lang VJ. Clinical reasoning education in the clerkship years: a cross-disciplinary national needs assessment. *PLoS One*. 2022;17(8):273250.
10. Brearley AM, Rott KW, Le LJ. A biostatistical literacy course: teaching medical and public health professionals to read and interpret statistics in the published literature. *J Stat Data Sci Educ*. 2023;31(3):286–294.
11. Federation of State Medical Boards of the United States; National Board of Medical Examiners. USMLE Content Outline. *US Medical Licensing Exam*. 2022. [https://www.usmle.org/sites/default/files/2022-01/USMLE\\_Content\\_Outline\\_o.pdf](https://www.usmle.org/sites/default/files/2022-01/USMLE_Content_Outline_o.pdf).
12. National Board of Osteopathic Medical Examiners. COMLEX-USA Master Blueprint. *NBOME*. 2023. [https://www.nbome.org/wp-content/uploads/pdf/COMLEX-USA\\_Master\\_Blueprint\\_2023.2.pdf](https://www.nbome.org/wp-content/uploads/pdf/COMLEX-USA_Master_Blueprint_2023.2.pdf).
13. Bonasso P, Lucke-Wold B, Iii, Reed Z, Bozek J, Cottrell S. Investigating the impact of preparation strategies on USMLE Step 1 performance. *MedEdPublish*. 2015;4(1):5.
14. Finn E, Ayres F, Goldberg S, Hortsch M. Brave new e-world: medical students' preferences for and usage of electronic learning resources during two different phases of their education. *FASEB Bioadv*. 2022;4(5):298–308.
15. Narayanan R, Nugent R, Nugent K. An investigation of the variety and complexity of statistical methods used in current internal medicine literature. *South Med J*. 2015;108(10):629–634.
16. Macdougall M, Cameron HS, Maxwell S. Medical graduate views on statistical learning needs for clinical practice: a comprehensive survey. *BMC Med Educ*. 2019;20(1):1.
17. Oster RA, Devick KL, Thurston SW. Learning gaps among statistical competencies for clinical and translational science learners. *J Clin Transl Sci*. 2020;5(1):12.
18. Yi D, Ma D, Li G. Statistical use in clinical studies: is there evidence of a methodological shift?. *PLoS One*. 2015;10(10):140159.
19. Cvetanovich GL, Fillingham YA, Harris JD, Erickson BJ, Verma NN, Bach BR. Publication and level of evidence trends in the American Journal of Sports Medicine from 1996 to 2011. *Am J Sports Med*. 2015;43(1):220–225.
20. Tyagi A, Garg D, Mohan A. Overview of statistical methods usage in Indian anaesthesia publications. *Indian J Anaesth*. 2022;66(11):783–788.
21. Alexander BK, Paul KD, Solar S. How has statistical testing in orthopedics changed over time? an assessment of high impact journals over 25 years. *J Surg Educ*. 2023;80(7):52.
22. Arnold LD, Braganza M, Salih R, Colditz GA. Statistical trends in the Journal of the American Medical Association and implications for training across the continuum of medical education. *PLoS One*. 2013;8(10):77301.
23. Araoye I, He JK, Gilchrist S, Stubbs T, MCGWIN G, Ponce BA. Collaborative orthopaedic educational research group. a national survey of orthopaedic residents identifies deficiencies in the understanding of medical statistics. *J Bone Joint Surg Am*. 2020;102(5):19.
24. Alzahrani SH, Al-Khail A, A B. Resident physician's knowledge and attitudes toward biostatistics and research methods concepts. *Saudi Med J*. 2015;36(10):240.
25. Zhelev Z, Garside R, Hyde C. A qualitative study into the difficulties experienced by healthcare decision makers when reading a Cochrane diagnostic test accuracy review. *Syst Rev*. 2013;2(1):32.
26. Susarla SM, Lifchez SD, Losee J, Hultman CS, Redett RJ. Plastic surgery residents' understanding and attitudes toward biostatistics: a national survey. *Ann Plast Surg*. 2016;77(2):231–236.
27. Couture F, Nguyen DD, Bhojani N, Lee JY, Richard PO. Knowledge and confidence level of Canadian urology residents toward biostatistics: a national survey. *Can Urol Assoc J*. 2020;14(10):E514–E519.
28. Whiting PF, Davenport C, Jameson C. How well do health professionals interpret diagnostic information? a systematic review. *BMJ Open*. 2015;5(7):8155.
29. Naeger DM, Kohi MP, Webb EM, Phelps A, Ordovas KG, Newman TB. Correctly using sensitivity, specificity, and predictive values in clinical practice: how to avoid three common pitfalls. *A J Roentgenology*. 2013;200(6).
30. Ghosh AK, Ghosh K, Erwin PJ. Do medical students and physicians understand probability?. *QJM*. 2004;97(1):53–55.
31. Najmi A, Sadasivam B, Ray A. How to choose and interpret a statistical test? an update for budding researchers. *J Family Med Prim Care*. 2021;10(8):767.
32. Enders FT, Lindsell CJ, Welty LJ. Statistical competencies for medical research learners: what is fundamental?. *J Clin Transl Sci*. 2017;1(3):146–152.
33. Ilic D, Maloney S. Methods of teaching medical trainees evidence-based medicine: a systematic review. *Med Educ*. 2014;48(2):124–135.
34. Buchanan BB, Allen GB, Pumphrey CM, Swaiti AR, Harris HM, Boyer PJ. Engaging medical students in the foundational curriculum using third-party resources. *Med Teach*. 2022;45(3):336.