BMC Medical Education

RESEARCH

Open Access

Imbalanced specialty representation of USMLE and NBME test writers

Mary T. Mahoney[1*], Lauren C. Linkowski[2], Michael D. Mix[3], Roger Wong[4,5], Austin J. Sim[6], Sarah E. Hoffe[7] and Karna T. Sura[3]

Abstract

Purpose The United States Medical Licensing Examination (USMLE) is an examination series required for allopathic physician licensure in the United States (US). USMLE content is created and maintained by the National Board of Medical Examinations (NBME). The specialty composition of the USMLE and NBME taskforce members involved in the creation of examination content is currently unknown.

Methods Using the 2021 USMLE and 2021 NBME Committees and Task Forces documents, we determined each member's board-certified primary specialty and involvement in test material development committees who we dubbed "test writers". Total active physicians by primary specialty were recorded from the 2020 Physician Specialty Data Report published by the Association of American Medical Colleges (AAMC). Descriptive statistics and chi-square analysis were used to analyze the cohorts.

Results The USMLE and NBME test writer primary specialty composition was found to be significantly different compared to the US active physician population (USMLE $\chi^2[32] = 172$, $p < .001$ and NBME $\chi^2[32] = 200$, $p < .001$). Only nineteen specialties were represented within USMLE test writers, with three specialties being proportionally represented. Two specialties were represented within NBME test writers. Obstetrics and Gynecology physicians were proportionally represented in USMLE but not within NBME test writers. Internal Medicine (IM) accounts for the largest percentage of all USMLE test writers (60/197, 30%) with an excess representation of 31 individuals.

Conclusions There is an imbalance in the specialty representation of USMLE and NBME test writers compared to the US active physician population. These findings may have implications for the unbiased and accurate portrayal of topics in such national examinations; thus, future investigation is warranted.

Keywords Undergraduate Medical Education, United States medical licensing Exam® (USMLE®), Medical Specialities, Standardized medical examinations

*Correspondence:
Mary T. Mahoney
marymahoneymd02@gmail.com
[1]Department of Radiation Oncology, Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia, PA, USA
[2]Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA, USA
[3]Department of Radiation Oncology, SUNY Upstate Medical University, Syracuse, NY, USA
[4]Department of Public Health and Preventive Medicine, SUNY Upstate Medical University, Syracuse, NY, USA
[5]Department of Geriatrics, SUNY Upstate Medical University, Syracuse, NY, USA
[6]Department of Radiation Oncology, James Cancer Hospital, The Ohio State University Comprehensive Cancer Center, Columbus, OH, USA
[7]Department of Radiation Oncology, Moffitt Cancer Center, Tampa, FL, USA

Introduction

The United States Medical Licensing Examination (USMLE) is an examination series required for allopathic physician licensure in the United States (US) and for international medical graduates wishing to practice in the US [1]. The USMLE series is composed of the Step 1, Step 2 Clinical Knowledge (CK), and Step 3 examinations which are landmark assessments throughout the continuum of medical training of basic science, clinical science, and internship [1]. USMLE content is, in part, maintained by the National Board of Medical Examinations (NBME) [1, 2]. The National Board of Medical Examiners (NBME) also creates the Clinical Science Subject Exam ('shelf exams'), which are standardized exit examinations of the core clinical clerkship rotations for many US medical students [3].

Given the universality and importance of the USMLE as an assessment in the allopathic medical training coupled with breadth of topics covered on the examination [4], careful evaluation of the members responsible for the examination questions and content is indicated. USMLE test questions are derived from various USMLE test committees which, per the USMLE, represent a "national faculty of medicine" [1]. Likewise, the NBME claims that the test development committees responsible for subject examinations are composed of "health professionals who are experts in the area" [2]. To evaluate these claims, we compared these test writer populations to the US actively practicing physician population.

Methods

The study was exempt by the SUNY Upstate Medical University Institutional Review Board. Listed individuals from USMLE Test Material Development Committees of the 2021 USMLE Committees and Task Forces [5] and two NBME subject examination task forces of "Family Medicine", and "Obstetrics and Gynecology" from the 2021 NBME Committees and Task Forces [6] were included. A general web search on the individual was conducted for a board-certified specialty, and if one existed, the individual was added to the "USMLE Test Writers" or "NBME Test Writers" cohort, respectively. Individuals already consented to be listed on the directories and furthermore were deidentified during our study. As such, no consent to participate declarations were collected.

In order to conduct a chi-square goodness of fit analysis, a control population was constructed from the 2020 Physician Specialty Data Report [7] published by Association of American Medical Colleges (AAMC), containing the number of total active physicians by predefined 47 unique specialties. To create a control population that is congruent in primary specialty division, and thus have a meaningful comparison to the USMLE and NBME test writers, we made modifications to create our general physician population. We combined all pediatrics subspecialties ("Pediatrics Anesthesiology, "Pediatrics Cardiology", "Pediatrics Critical Care Medicine", "Neonatal-Perinatal Medicine", "Pediatrics Hematology/Oncology") and counted them all as "pediatrics" specialty. Similarly, "Child and Adolescent Psychiatry" was combined with, and counted under "Psychiatry". "Pain medicine/Pain Management" and "Sports Medicine" were omitted given unclear alternative primary specialty placement. Because no urologist appeared in the USMLE and NBME taskforce, urologists were likewise removed from the control population. This generated a total of 33 unique primary specialties. This control population was dubbed "Modified Active US Physicians by Primary Specialty" (Supplemental Table 1).

Two separate chi-square goodness of fit tests were conducted to examine whether the observed proportions by specialty for USMLE and NBME test writers differed from expected proportions based on active physician specialties. Analyses were performed in Stata 18 (StataCorp LLC, College Station, TX, US) with two-tailed tests at a 0.05 significance level. Due to small expected frequencies ($N<5$) in several cells from the chi-square, we also performed sensitivity analyses through exact multinomial goodness of fit test (Monte Carlo method) using the Stata *mgof* module [8].

Results

There were 197 USMLE and 25 NBME test writers in 2021. There was no overlap of individuals between these two test writer cohorts. The USMLE test writer specialty composition was significantly different compared to the specialty composition in the US active physician population ($\chi^2[32]=172$, $p<.001$). The NBME test writer specialty composition was also significantly different compared to that of the US actively practicing physician population ($\chi^2[32]=200$, $p<.001$). Sensitivity analyses using exact tests for small samples also generated statistically significant differences ($p<.001$) for the USMLE and NBME test writers, indicating our findings are robust. A summary of US actively practicing physicians by specialty compared to USMLE and NBME test writer representation can be found in Tables 1 and 2, respectively.

Nineteen specialties were represented within USMLE test writers, with three specialties including neurology, physical medicine and rehabilitation, and neurosurgery were found to be proportionally represented. Despite only accounting for 14.8% of actively practicing physicians, Internal Medicine physicians account for the largest percentage of USMLE (29.1%) with a representation with an excess of 31 USMLE test writers.

Only two specialties were represented in NBME test writers within our scope of analysis of shelf examinations:

Table 1 Primary Specialty representation of USMLE Test writers. Captures the relative representation of the 33 unique specialties among USMLE test writers and modified active US practicing physician population. The difference column is denoted with + for specialties in which an excess of members is present and denoted - for specialties lacking proportional representation

Board Certified Specialty	Modified Total Active Physicians (%)	Observed on USMLE Task Taskforces and Committees (%)	Observed Representation USMLE Task Forces and Committees (*n*)	Expected Representation on USMLE Task Force and Committees (*n*)	Difference in Representation (Observed- Expected) (*n*)
Internal Medicine (Hospital Medicine and Primary Care only)	14.8	30.5	60	29	+31
Family Medicine	14.6	12.2	24	29	-5
General Surgery	3.1	13.2	26	6	+20
Infectious Disease	1.2	2.0	4	2	+2
Pediatrics	9.6	10.2	20	19	+1
OBGYN	5.3	5.6	11	10	+1
Geriatrics	0.7	0.0	0	1	-1
Critical Care Medicine	1.6	0.0	0	3	-3
Emergency Medicine	5.6	6.1	12	11	+1
Pathology	1.6	1.0	2	3	-1
Psychiatry	6.0	6.6	13	12	+1
Neurology	1.7	1.5	3	3	0
Anesthesiology	5.2	2.0	4	10	-6
Ophthalmology	2.4	2.0	4	5	-1
Plastic Surgery	0.9	0.5	1	2	-1
Allergy and Immunology	0.6	0.0	0	1	-1
Cardiology	3.3	0.0	0	7	-7
ENT	1.2	0.5	1	2	-1
Pulmonology	0.6	0.0	0	1	-1
Metabolism/ Endocrinology	1.0	0.0	0	2	-2
Dermatology	1.5	0.5	1	3	-2
Physical Medicine and Rehabilitation	1.2	1.0	2	2	0
Hematology and Oncology	2.0	0.0	0	4	-4
Neurosurgery	0.7	0.5	1	1	0
Radiology (Interventional and Diagnostic)	4.4	0.5	1	9	-8
Gastroenterology	1.9	0.0	0	4	-4
Rheumatology	0.8	0.0	0	2	-2
Nephrology	1.4	0.0	0	3	-3
Vascular Surgery	0.5	0.0	0	1	-1
Cardiothoracic Surgery	0.6	0.0	0	1	-1
Orthopedic Surgery	2.7	0.0	0	5	-5
Radiation Oncology	0.7	0.0	0	1	-1
Internal Medicine - Pediatrics	0.7	3.6	7	1	+6

Family Medicine (9/25, 36%) and Obstetrics and Gynecology (16/25, 64%). Family medicine physicians account for 14.8% while Obstetrics and Gynecology physicians account for 5.3% of all actively practicing US physicians.

Discussion

To our knowledge, this is the first study to investigate the primary specialty composition of the USMLE and NBME members involved in the creation of examination content. We found that USMLE and NBME test writers' primary specialty composition did not reflect the US actively practicing physicians. This suggests certain medical specialties may not be appropriately represented while writing standardized medical examination content. Potential consequences include introducing bias, imbalance, and possible inaccuracies in standardized test questions [9, 10]. When reviewing USMLE examination questions regarding obesity, the Obesity Society of America and American Board of Obesity Medicine cited a lack of obesity experts as a contributing to the detected bias in such questions [9]. Critical care topics accounted for 19% of all sampled USMLE Step 2 questions [10], despite our analysis citing 1.6% of US active practicing physicians in the specialty. Although "critical care" could potentially encompass multiple specialties and subject topics, there

Table 2 Primary Specialty representation of NBME Test writers. Captures the relative representation of the 33 unique specialties among NBME test writers and modified active US practicing physician population. The difference column is denoted with + for specialties in which an excess of members is present and denoted - for specialties lacking proportional representation

Board Certified Specialty	Modified Total Active Physicians (%)	Observed on NBME Test Writers (%)	Observed Representation on NBME Test Writers (people)	Expected Representation on NBME Test Writers (people)	Difference in Representation (Observed- Expected) (people)
Internal Medicine (Hospital Medicine and Primary Care only)	14.8	0	0	4	-4
Family Medicine	14.6	36	9	3	6
General Surgery	3.1	0	0	1	-1
Infectious Disease	1.2	0	0	0	0
Pediatrics	9.6	0	0	2	-2
OBGYN	5.3	64	16	1	15
Geriatrics	0.7	0	0	0	0
Critical Care Medicine	1.6	0	0	0	0
Emergency Medicine	5.6	0	0	1	-1
Pathology	1.6	0	0	0	0
Psychiatry	6.0	0	0	1	-1
Neurology	1.7	0	0	0	0
Anesthesiology	5.2	0	0	1	-1
Ophthalmology	2.4	0	0	1	-1
Plastic Surgery	0.9	0	0	0	0
Allergy and Immunology	0.6	0	0	0	0
Cardiology	3.3	0	0	1	-1
ENT	1.2	0	0	0	0
Pulmonology	0.6	0	0	0	0
Metabolism/ Endocrinology	1.0	0	0	0	0
Dermatology	1.5	0	0	0	0
Physical Medicine and Rehabilitation	1.2	0	0	0	0
Hematology and Oncology	2.0	0	0	0	0
Neurosurgery	0.7	0	0	0	0
Radiology (Interventional and Diagnostic)	4.4	0	0	1	-1
Gastroenterology	1.9	0	0	0	0
Rheumatology	0.8	0	0	0	0
Nephrology	1.4	0	0	0	0
Vascular Surgery	0.5	0	0	0	0
Cardiothoracic Surgery	0.6	0.0	0	0	0
Orthopedic Surgery	2.7	0.0	0	1	-1
Radiation Oncology	0.7	0.0	0	1	0
Internal Medicine -Pediatrics	0.7	0.0	0	1	0

could still possibly be an imbalanced content-to-writer ratio.

Our study also raises the question regarding the size of a specialty as a justification for exclusion from aspects of general medical examination creation. For example, cancer is found throughout the USMLE content outline [4]. Oncology treatments are often multidisciplinary, to which Radiation Therapy serves a critical role. A commonly provided reason for the lack of Radiation Oncology (RO) representation in undergraduate medical education and medical licensing examinations is that RO is too small of a field to justify inclusion. Our findings do demonstrate that Radiation Oncologists are a small proportion of active US physicians (0.7%). However, many of other fields of medicine are comparable, compromising < 1% of the active physician population: Geriatric (0.7%), Medicine-Pediatrics (0.7%), Allergy and Immunology (0.6%), Pulmonology (0.6%), Neurological Surgery (0.7%), Rheumatology (0.8%), Vascular Surgery (0.5%), and Thoracic Surgery (0.5%). For comparable-sized specialties of 0.7%, only neurology surgery was equally underrepresented in the USMLE test writers. When investigating the USMLE taskforce committees [5] directly, only one Radiation Oncologist was identified and present on the Step 1 Standard Setting Committee. These findings suggest that Radiation Oncology expert

review may be absent from USMLE Step 2CK and Step 3 questions. This raises concern regarding important topics in oncology that need to be covered for all providers and should be further investigated.

One of the core missions of the USMLE "to develop and improve assessment for licensure with intent of assessing physicians more accurately and comprehensively" [1]. Our findings demonstrate a broad imbalance of specialty representation, often regardless of specialty size, which highlights an area for improvement. When NBME was called to integrate nutritional content into Step 1 and Step 2, they explicitly included a nutritional task force composed of multidisciplinary experts in the field when constructing the new examination questions [11]. It would not be unreasonable to reformat current task forces and properly balance expert opinion of tested topics.

It should be noted that NBME subject examinations are more specific to the specialty tested unlike the USMLE Step examination series, which reflects a general medical examination. As such, these NBME subject examinations may not require as rigorous balancing since the tested content may usually reflects a specific specialty. This may explain our findings of only Family Medicine and OBGYN specialty representation on the 2021 NBME Taskforce committee directories. However, there is a previous precedent to include a multidisciplinary panel for creations of tests questions [11]; thus, NBME could have a similar approach to review questions to confirm little to no bias is present in the line of questions.

A limitation of the study is the reliance on the 2021 USMLE and NBME Taskforce and Committees directories [5, 6] as a single reference point. Future studies comparing the taskforce committee compositions longitudinally are needed to determine if this specialty imbalance is a systemic finding. It should also be noted possible error regarding calculations of comparing observed vs. expected test writers given the small counts of individuals within the committees and smaller specialties. Board certification status may also not accurately reflect true specialty practice. For example, if a family medicine or internal medicine physician is boarded in geriatrics but primarily practices outside the field, are they still considered geriatricians? The authors of this paper did not contact the NBME directly on reading assistance on how to accurately assess the number and type of specialties on the various test materials development committees. Likewise, this study did not investigate any official USMLE or NBME examination content, nor the knowledge acquisition of students from test preparation. More studies are necessary to determine if there are biased portrayals of topics within these examinations, particularly in the underrepresented fields of medicine within medical education.

Conclusions

From our analysis, the primary specialty composition of USMLE and NBME test writers is not reflective of that of the active practicing physician population. Our findings demonstrate a broad imbalance of specialty representation, often regardless of specialty size, which highlights an area for improvement. These findings may have implications on the unbiased and accurate portrayal of topics in such national examinations. More studies are necessary to determine if there are biased portrayals of topics within these examinations, particularly in the underrepresented fields of medicine within medical education.

Abbreviations

USMLE® United States Medical Licensing Exam®
NBME© National Board of Medical Examiners

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12909-024-05976-0.

Supplementary Material 1

Acknowledgements

We would like to thank the Radiation Oncology Education Collaboration Study Group (ROECSG) Undergraduate Medical Education Working Group, particularly Dr. Jillian R Gunther MD PhD and Dr. Steve E Braunstein MD PhD, for their support with this project.

Author contributions

Concept and design: All authors. Acquisition, analysis, or interpretation of data: M.T.M., L.C.L., R.W. Drafting of the manuscript: M.T.M., K.T.S., L.C.L. Preparation of Table 1, and 2: M.T.M., K.T.S. Critical review of the manuscript for important intellectual content: All authors. Statistical analysis: R.W.Administrative, technical, or material support: M.T.M. Supervision: M.D.M, A.J.S., S.E.H., K.T.S.

Funding

The authors do not have any sources of support, including grants, funding, or other assistance, whether from public or private sources.

Data availability

The coded data that support the findings of this study are available on request from the corresponding author [MTM]. The data are not publicly available due to privacy or ethical restrictions.

Declarations

Ethical approval

This study was found to be exempt from IRB review by SUNY Upstate IRB.

Human ethics and consent to participate declarations

Individuals already consented to be listed on the directories and furthermore were deidentified during our study. As such, no consent to participate declarations were collected.

Previous presentations

The authors report that selected findings of this work were presented during the American Radium Society® 105th Annual Meeting on 5/19/2023 in Lahaina, HI, USA and was subsequently electronically published in the Proceedings of the American Radium Society® 105th Annual Meeting in *American Journal of Clinical Oncology.*

Data

This study was based on the 2021 USMLE Taskforce and Committee Directory, 2021 NBME Taskforce and Committee Directory, and the AAMC 2020 Physician

Specialty Data Report, which are all publicly available documents as cited in the references.

Conflict of interest
The authors have no competing interests to declare that are relevant to the content of this article.

Received: 19 April 2024 / Accepted: 30 August 2024
Published online: 09 September 2024

References

1. Federation of State Medical Boards of the United States Inc. (FSMB) and National Board of Medical Examiners (NBME). About the USMLE. 2023. https://www.usmle.org/about-usmle. Accessed January 2, 2023.
2. National Board of Medical Examiners. Item Co-Creation. 2023. https://www.nbme.org/about-nbme/our-collaborations/item-co-creation. Accessed August 16 2023.
3. National Board of Medical Examiners. Clinical Science Subject Examinations. 2023. https://www.nbme.org/assessment-products/assess-learn/subject-exams/clinical-science. Accessed January 2, 2023.
4. Federation of State Medical Boards of the United States Inc. (FSMB) and National Board of Medical Examiners (NBME). USMLE® Content Outline. 2022. https://www.usmle.org/prepare-your-exam/step-1-materials/step-1-content-outline-and-specifications. Accessed July 1 2022.
5. National Board of Medical Examiners. 2021 USMLE Taskforce and Committee Directory. https://www.nbme.org/about-nbme/our-collaborations/item-co-creation. 2022. Accessed September 6, 2022.
6. 2021 NBME National Board of Medical Examiners, Taskforce, and Committees. 2022. https://www.nbme.org/about-nbme/our-collaborations/item-co-creation. Accessed October 1, 2022.
7. 2020 Physician Specialty Data Report. AAMC. 2021. https://www.aamc.org/data-reports/workforce/data/active-physicians-largest-specialties-2019. Accessed October 15, 2022.
8. Jann BMGOF. Stata Module to perform goodness-of-fit tests for multinomial data. Statistical Software Components. January 18, 2021. https://ideas.repec.org/c/boc/bocode/s456854.html. Accessed July 10, 2024.
9. Kushner RF, Butsch WS, Kahan S, Machineni S, Cook S, Aronne LJ. Obesity Coverage on Medical Licensing Examinations in the United States. What Is Being Tested? Teaching and Learning in Medicine. 2017. https://doi.org/10.1080/10401334.2016.1250641
10. Frankel HL, Rogers PL, Gandhi RR, Freid EB, Kirton OC, Murray MJ. What is taught, what is tested: findings and competency-based recommendations of the undergraduate medical education committee of the Society of Critical Care Medicine. Crit Care Med. 2004;32:1949–56.
11. Hark LA, Iwamoto C, Melnick DE, Young EA, Morgan SL, Kushner RF, et al. Nutrition coverage on medical licensing examinations in the United States. Am J Clin Nutr. 1997;65:568–71.

Publisher's note