Research supporting the validity of licensing decisions made by state medical boards based upon USMLE

Since its introduction in 1992, the United States Medical Licensing Examination® (USMLE®) has provided the medical licensing community with a high-quality assessment tool ensuring that licensed physicians are held to a rigorous and reliable standard. Passing USMLE Step 1, 2 Clinical Knowledge (CK), and 3 ensures that physicians have demonstrated they understand and can apply medical knowledge and clinical reasoning to provide safe and effective patient care.

USMLE has administered more than three million Step administrations since 1992—providing much of the groundwork for an extensive research program that continues to produce numerous peer-reviewed contributions to the professional literature of medical licensing, education and training.

Staff at NBME and FSMB working on the USMLE program have assembled a collection of key recent articles evaluating the predictive validity of USMLE. Predictive validity involves studying how the exam scores are associated with relevant future outcomes and thus can provide strong evidence that the exam truly measures competencies related to safe and effective practice. Extensive research has focused on USMLE performance and its correlation with other key measures, such as performance on other professional assessments, residency outcomes, disciplinary actions by state medical boards, and, most importantly, patient outcomes.

These studies augment the validity evidence from the rigorous exam development, scoring, and standard setting practices that collectively support the validity of licensing decisions informed, in part, by requiring successful completion of USMLE Steps 1, 2 and 3. These key articles are shared in their entirety or with citation and brief summary. The attached appendix provides links to the online print versions.

Our intent is to share relevant literature supporting medical boards' continued utilization of the USMLE as the primary assessment tool in the decision to issue a full, unrestricted medical license. Questions about these studies or interpretations of this data may be directed to Daniel Jurich, PhD, NBME Associate Vice President for USMLE.

David Johnson, MA
FSMB Chief Assessment Officer

Alex J. Mechaber, MD, MBA
NBME Vice President for USMLE

Research Demonstrating Predictive Validity for USMLE

There are four categories of research are presented here.
- Studies correlating performance on USMLE with direct patient outcomes.
- Studies correlating performance on USMLE with disciplinary actions taken by state medical boards.
- Studies correlating performance on USMLE with performance in residency training.
- Studies correlating performance on USMLE with other professional assessments such as outcomes on specialty board certification exams.

<u>USMLE and patient outcomes</u>
Typically regarded as the most important form of predictive validity research in medical assessment is that which explores the association with patient outcomes. Such research can be difficult to construct and execute given the statistical complexity of connecting exam performance to future patient outcomes. Yet, this evidence remains the "gold standard" for predictive validity research. A pair of studies from 2024[1] and 2014[2] spotlight USMLE performance and patient outcomes involving treatment for common inpatient diagnoses, length of in-hospital stay and in-hospital mortality. Both studies identified a correlation between USMLE performance and improved patient outcomes in the specified areas after accounting for various other relevant factors. See the two studies attached here.

<u>USMLE and state board disciplinary actions</u>
Multiple studies have explored the association between USMLE performance and subsequent likelihood of disciplinary action by a state medical board[3,4]. The findings indicate that higher performance on each USMLE examination relates to a lower likelihood of disciplinary action. Another study exploring this outcome revealed that more attempts on USMLE Step exams were associated with an increased likelihood of subsequent disciplinary action[5].

<u>USMLE and performance in residency training</u>
The introduction of competency "milestones" by the Accreditation Council for Graduate Medical Education (ACGME) into residency training has provided another point for comparison with USMLE performance. It is not unreasonable to expect performance measures and milestones to align with competencies assessed on USMLE. The attached 2021 study provides one example of this, showing

incremental validity evidence for support of Step 1 and 2 scores with emergency medicine milestones[6].

USMLE and other professional examinations

One of the most common forms of predictive validity evidence explores the relationship between an exam and other assessments within that field. Thus, another line of research has explored the relationship between USMLE and other medical education examinations, such as board certification exams. These studies typically use USMLE as a control to account for prior knowledge before the specialties in-training examination. However, the studies often reveal that USMLE performance also strongly relates to the certification exam.

For example, a 2020 study published in *Academic Medicine* explored the correlation between performance on USMLE Step exams, the in-training exam of the American College of Physicians, and the American Board of Internal Medicine's (ABIM) certifying examination[7]. This study showed that while no individual USMLE Step score was as strongly predictive of the ABIM certifying exam score as the internal medicine in-training exam score, the combined relative contribution of all three USMLE Step scores was similar to that of the in-training score.

Comparable research has been done and published for numerous other specialties, including: infectious disease (2015)[8], adult rheumatology (2015)[9], hematology and medical oncology (2016)[10], cardiology (2017)[11] and nephrology (2018)[12].

The research spotlighted here shows a consistent positive relationship between USMLE performance and key external outcomes—precisely what one should expect of a high-stakes examination for medical licensure. The studies shared and cited here provide a compelling basis for the continued validity of the licensing decisions based, in part, on the successful completion of the USMLE Step sequence.

Appendix: References

Associations with Patient Outcomes

1. [Higher performance on USMLE sequence associated with improved patient outcomes (decreased mortality and length of stay) for Pennsylvania hospitalizations](#) (2024)
2. [For International graduates', higher Step 2 CK scores correlated with improved mortality outcomes](#) (2014)

Associations with Fewer State Board Disciplinary Sanctions/Actions

3. [Higher Step 1 and Step 2 CK performance correlated with lower sanctions](#) (2017)
4. [Higher Step 3 performance correlated with lower sanctions](#) (2022)
5. [Repeat USMLE attempts associated with a higher probability of disciplinary actions](#) (2021)

Associations with Residency Outcomes

6. [Small but statistically significant relationships between USMLE and expected ACGME Emergency Medicine residency milestones](#) (2020)

Associations with Board Certification Examinations

A sample of the many articles within this program of research**:**

7. [Internal Medicine Certification](#) (2020)
8. [Infectious Disease Certification](#) (2015)
9. [Adult Rheumatology Certification](#) (2015)
10. [Hematology and Medical Oncology Certification](#) (2016)
11. [Cardiology Certification](#) (2017)
12. [Nephrology Certification](#) (2018)