BCM Medical Student Review

Issue 2
Fall 2026

Editorial Board

Peer Reviewers

Aryan Jain
Chad Mourino
Rohan Phadke
Lauren Reppert
Nisha Saif
Samer Salman
Malay Shah
Sarah Soubra
Geoffrey Zhang

Undergraduate Liaison

Salma Gaballa

On the Second Issue

Dear Readers,

It is our pleasure to welcome you to the second issue of the *BCM Medical Student Review* (*BMSR*).

Established by medical students at Baylor College of Medicine, the *BMSR* was created to cultivate scholarship, encourage intellectual curiosity, and provide a platform where trainees can engage in meaningful academic discourse. We are proud to feature original work from contributors across the country, including undergraduate students, medical students, and physicians, whose diverse perspectives enrich conversations across the breadth of medicine. Through scientific scholarship, thoughtful commentary, and creative expression, this issue reflects our continued commitment to amplifying trainee voices and fostering collaboration across all stages of medical education.

The publication of this second issue also marks an exciting period of growth for the journal. As we continue to expand our mission, we are strengthening our engagement with undergraduate students through the addition of an Undergraduate Liaison, creating new opportunities for mentorship, collaboration, and early involvement in academic publishing.

We are also pleased to announce our collaboration with Digital Commons@TMC, the open-access institutional repository supported by the distinguished Texas Medical Center Library. This partnership represents an important step in increasing the accessibility and long-term preservation of the scholarship published in the *BMSR*. As we continue this transition, we are pursuing an International Standard Serial Number (ISSN) and implementing permanent Digital Object Identifier (DOI) links for published articles. Together, these initiatives will enhance the discoverability, citation, and global visibility of our authors' work, allowing readers around the world to access and engage with the scholarship published in our journal. Permanent DOI assignment will also provide each article with a persistent digital identity, ensuring reliable access and facilitating future academic citation.

None of these accomplishments would be possible without the dedication of our editorial board, peer reviewers, faculty advisors, and authors. We are deeply grateful to everyone who devoted their time, expertise, and enthusiasm to making this issue possible. The thoughtful scholarship submitted by our contributors and the careful work of our reviewers continue to elevate the quality and impact of the journal.

As the *BMSR* continues to evolve, we remain committed to promoting rigorous scholarship, encouraging interdisciplinary dialogue, and creating opportunities for trainees to develop as authors, reviewers, and future leaders in academic medicine. We are excited to continue building this community and look forward to the future of the journal.

Thank you for reading, and we hope you enjoy this second issue.

Sincerely,

Ritu Sampige
Editor-in-Chief

Allen Kuncheria
Deputy Editor

Julia Morales
Deputy Editor

Nicholas O'Brien
Deputy Editor

Disclaimer: The views and opinions expressed in this publication are those of the individual authors and do not necessarily reflect the official policies or positions of Baylor College of Medicine.

Contents

Issue 2

ORIGINAL RESEARCH

Shifting Coverage, Shifting Risk: A 10-Year Snapshot of Health Insurance Trends and Hypertension Among U.S. Youth

Aydee Hinojosa, BS[1*], Deepali K. Ernest, PhD, MPH[2], Bipin Singh, PhD(c), MPH[2], Elizabeth A. Onugha, MD, MSCI[2,3], Shreela V. Sharma, PhD, RDN, LD[2], Jayna M. Dave, PhD[2,4]

[1]School of Medicine, Baylor College of Medicine, Houston, Texas, USA
[2]Center for Health Equity, Department of Epidemiology, UTHealth School of Public Health, Houston, Texas, USA
[3]Renal Section, Department of Pediatrics, Texas Children's Hospital, Baylor College of Medicine, Houston, Texas, USA
[4]USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA
[*]Corresponding author: Aydee Hinojosa, BS; Email: aydee.hinojosa@bcm.edu
Disclosures: This project was partially supported by the UTHealth Center for Health Equity. The authors declare no competing financial or non-financial conflicts of interest.
Funding: This work was funded, in part, by federal funds from the USDA/ARS under Cooperative Agreement no. 58-3092-5-008 (J.M.D.). The contents of this publication do not necessarily reflect the views or policies of the USDA, nor does the mention of trade names, commercial products, or organizations imply endorsement from the U.S. government.
Acknowledgements: The authors thank the UTHealth Center for Health Equity for administrative and research support.

ABSTRACT

Background: As pediatric hypertension becomes increasingly prevalent, effective policies must prioritize prevention, early detection, and coordinated care to reduce cardiovascular risk among youth. This study aims to examine trends in hypertension prevalence and health insurance coverage among U.S. adolescents aged 13–18 from 2013 to 2023, and to assess how changes in insurance patterns relate to shifts in hypertension rates.

Methods: Using serial cross-sectional NHANES data (2013–2023), we examined trends in hypertension and insurance coverage among 4,798 adolescents aged 13–18 years. Blood pressure was measured clinically; insurance and sociodemographics were self-reported. Weighted analyses accounted for the complex survey design.

Results: Hypertension prevalence more than doubled, from 1.9% in 2013–2014 to 4.6% in 2021–2023. Over the same period, the uninsured rate halved, private coverage declined, and enrollment in Medicaid/Medicare rose. Despite gains in public insurance, hypertension rates continued to rise over the study period. Adjusted analyses examining associations between insurance status and hypertension were not statistically significant.

Conclusion: Enrollment in public insurance has improved but has not curbed the growing hypertension burden in youth. These findings suggest that insurance coverage alone may not fully address the growing burden of hypertension among youth. Comprehensive prevention, screening, and coordinated care are needed to protect adolescent cardiovascular health.

Keywords: Adolescent Hypertension; Health Insurance; Healthcare Access; Social Determinants of Health; United States

Introduction

Hypertension affects approximately 5% of U.S. adolescents aged 13–18 years and is a critical early marker for future cardiometabolic conditions, including heart failure, chronic kidney disease, stroke, and cognitive decline in adulthood.[1] While behavioral and biological risk factors such as physical activity, diet, age, and genetics are well-established contributors to youth hypertension, structural and social determinants—particularly access to health insurance—remain under-examined yet highly consequential.[2] Health insurance coverage is a pivotal and modifiable determinant of health that influences timely screening, diagnosis, and treatment of hypertension during adolescence. However, persistent disparities in insurance access—exacerbated during the vulnerable transition from pediatric to adult care—may contribute to delayed diagnosis and suboptimal management, especially among minority populations.[3] Vulnerabilities during this transition may involve aging out of parental insurance coverage, disruptions in continuity of care, and changes in eligibility for public programs. To improve our understanding of recent insurance coverage and hypertension trajectories, we conducted a nationally representative, serial cross-sectional analysis using NHANES data from 2013–2023 that evaluates trends in hypertension, health insurance coverage, and types of coverage among a diverse sample of U.S. youth. The study's special focus on youth is driven by the need for a better understanding of how healthcare coverage relates to hypertension in younger populations, to improve public health and clinical efforts to address hypertension in the U.S.

Materials and Methods

This serial cross-sectional analysis utilized 2013-2023 NHANES data collected in 2-year cycles.[4] Special steps were taken by NHANES to counter the impact of COVID-19 disruptions to sampling during the 2017-2020 and 2021-2023 cycles. NHANES participants between the ages of 13 to 18 were included in the study sample (N = 4,798). Blood pressure measurements and anthropometrics were measured during physical examinations by NHANES staff.[5] Sociodemographic and health insurance status, type, and coverage duration were self-reported by participants on NHANES questionnaires.[5] For this study, hypertension was defined as having a systolic pressure ≥130 and/or a diastolic pressure ≥80.[6] Weighted cross-tabulations determined the prevalence of health insurance characteristics and hypertension from 2013 to 2023. Weighted linear regression and chi-square tests (Rao-Scott p-values) with adjustments for age, sex, and race/ethnicity were used to assess temporal trends in hypertension and health insurance characteristics from 2013 to 2023. Additionally, survey-weighted logistic regression analyses were conducted to examine associations between health insurance status and hypertension, adjusting for age, sex, and race/ethnicity. Missing observations were considered missing at random and excluded from analysis. The type I error (α) was set to 5% and $p < 0.05$ was considered significant. All analyses were conducted in SAS version 9.4.

Results

This analysis included a sample of 4,798 U.S. youth (2013-14: N = 1,086; 2015-16: N = 1,008; 2017-2020: N = 1,513; 2021-2023: N = 1,191) with a hypertension prevalence of 3.7%. On average, participants had a mean age of 15.4 years and a mean BMI of 24.4 kg/m^2. The sample was evenly split between males and females and was primarily comprised of individuals who identified as non-Hispanic White and were from a high socioeconomic background. All sociodemographic characteristics remained stable across data-cycles (2013 to 2023). Additional participant characteristics are provided in Table 1. The prevalence of hypertension increased almost 2-fold among U.S. youth from 2013-2014 (1.9%) to 2021-2023 (4.6%), with the peak prevalence observed during 2017-2020 (4.8%) (p = 0.03). Across the same time period, the proportion of youth without health insurance coverage decreased by nearly 50%, from 9.9% in 2013-2014 to 5.3% in 2021-2023 (p = 0.04). Notably, the proportion of youth covered by private insurance declined by 10% from 2013 to 2023, accompanied by increased enrollment in government or state-sponsored coverage (particularly Medicare and Medicaid), although this increase was not statistically significant. Trends in hypertension and health insurance coverage over time are shown in Figure 1. Percentages for insurance type represent distributions among all participants. In adjusted analyses, uninsured participants had lower odds of hypertension compared to insured participants, although this was not statistically significant (aOR = 0.70; 95% CI: 0.25-1.99; p = 0.49). Similarly, participants with public insurance had higher odds of hypertension compared to those with private insurance, though this association was not statistically significant after adjustment (aOR = 1.43; 95% CI: 0.70-2.90; p = 0.32).

		NHANES Survey Year				
	Overall N = 4,798	**2013-2014 N = 1086**	**2015-2016 N = 1008**	**2017-2020 N = 1513**	**2021-2023 N = 1191**	***p*-value**
Age	15.4(0.03)	15.6(0.07)	15.3(0.06)	15.4(0.04)	15.5(0.09)	0.15
BMI	24.4(0.15)	24.4(0.35)	24.2(0.34)	24.5(0.23)	24.4(0.30)	0.92
Sex						0.86
Males	2190(51.1)	525(50.0)	499(50.7)	731(51.2)	435(52.4)	
Females	2115(48.9)	530(50.0)	478(49.3)	690(48.8)	417(47.6)	
Race & Ethnicity						0.33
Non-Hispanic White	1292(51.0)	262(54.2)	272(55.1)	437(50.6)	321(44.4)	
Non-Hispanic Black	950(13.2)	257(14.1)	224(13.5)	357(13.70)	112(11.3)	
Hispanic/Latino	1302(23.5)	351(21.9)	327(22.2)	340(23.4)	284(26.6)	
Other	761(12.2)	185(9.8)	154(9.2)	287(12.3)	135(17.8)	
Socioeconomic Status						0.45
Below poverty line	1108(19.8)	336(24.3)	239(17.0)	339(18.2)	194(20.7)	
Low	628(12.6)	147(12.1)	173(15.0)	210(12.3)	98(11.1)	
Middle	443(11.0)	104(11.2)	103(9.9)	155(11.2)	81(11.7)	
High	1692(56.6)	381(52.4)	372(58.0)	561(58.3)	378(56.6)	

[a] Mean(SD) reported; [b] N(%) reported; [c] tests whether the characteristic of interest differs significantly across NHANES data cycles.

Table 1. Distribution of Participant Characteristics Across NHANES Data Cycles (2013-2023)

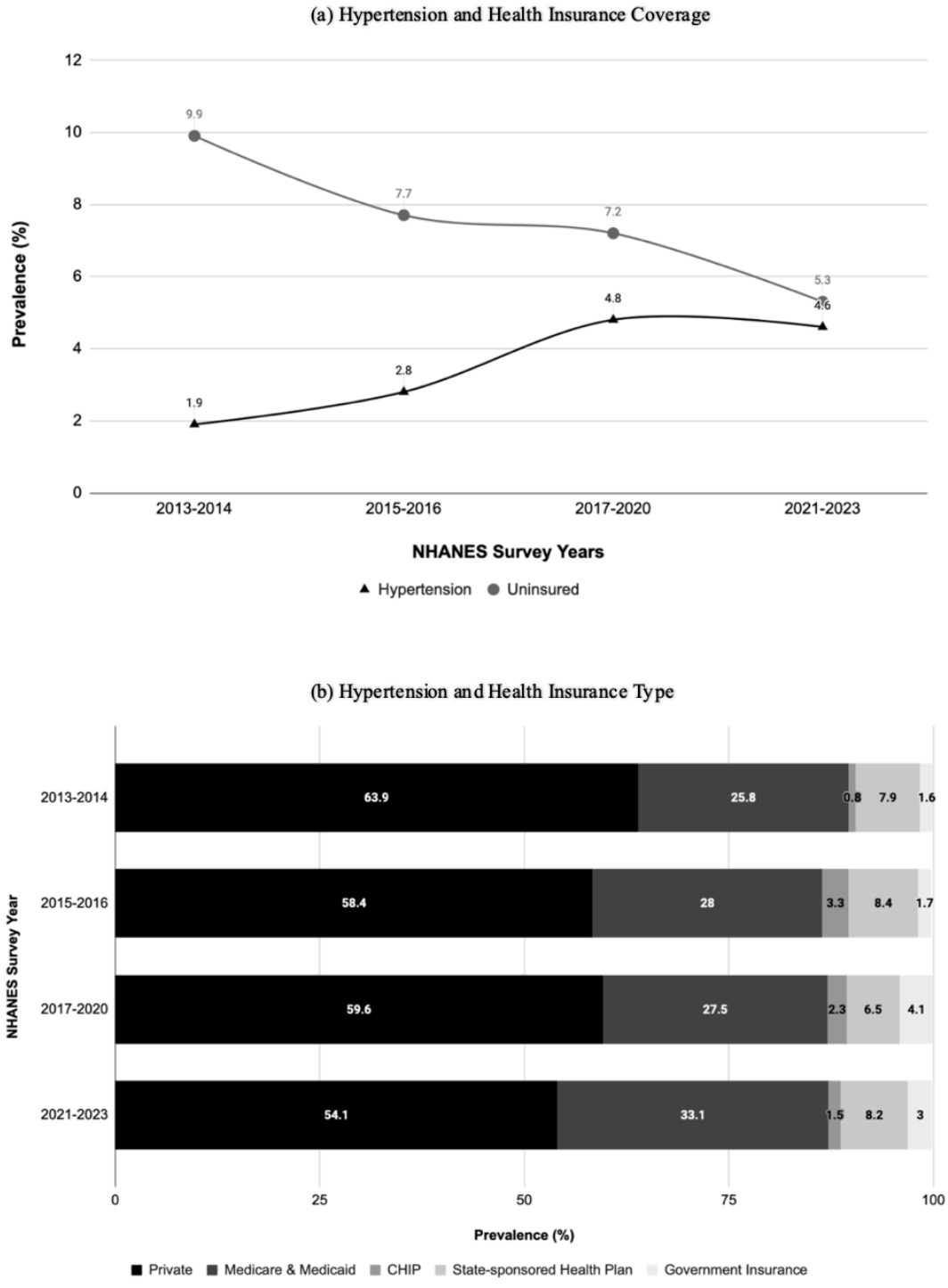

Figure 1. Prevalence Trends in (a) Hypertension Prevalence and Health Insurance Coverage, and (b) Distribution of Health Insurance Type Among Adolescents by NHANES Cycle. Percentages represent distributions among all participants.

Discussion

The sharp increase in hypertension prevalence among U.S. youth—from 1.9% in 2013-2014 to 4.6% in 2021-2023—signals an emerging pediatric cardiometabolic crisis. Adolescence is a critical window for shaping long-term health trajectories, and elevated blood pressure in youth is closely linked to early vascular changes, such as increased left ventricular mass and carotid intima-media thickness, which predict adult cardiovascular events.[7] Longitudinal data have confirmed that adolescent hypertension often persists into adulthood, raising lifetime risks of stroke, kidney disease, and premature mortality.[7,8] Compounding these clinical concerns, elevated blood pressure in youth has also been associated with neurocognitive deficits–even at a subclinical level–highlighting its broad developmental impact.[9] Although early intervention can substantially alter this trajectory, disparities in healthcare access, particularly during the transition from pediatric to adult care, continue to delay diagnosis and management in vulnerable populations.[7,9]

In this context, the nearly 50% reduction in uninsured youth over the last decade reflects a major public health achievement, driven largely by expanded enrollment in Medicaid, Medicare, and other state-sponsored programs. Medicaid now covers nearly 40% of U.S. children and over half of those with special healthcare needs. Its mandated Early and Periodic Screening, Diagnostic, and Treatment (EPSDT) benefit is particularly critical for identifying conditions like hypertension that may otherwise remain silent until irreversible damage has occurred. The increase in public insurance enrollment during this study period likely reflects the broader effects of the Affordable Care Act's Medicaid expansion and enhanced outreach via CHIP, which aimed to reduce coverage gaps in vulnerable populations in the U.S. It may also reflect an increase in poverty rates that further drives poorer dietary behaviors and thus higher obesity and hypertension rates. Meanwhile, the 10% decline in private insurance signals a shifting reliance towards public programs to meet the healthcare needs of youth, particularly among socioeconomically disadvantaged households.

Yet, the rise in hypertension prevalence despite expanded coverage warrants further investigation into factors contributing to these concurrent trends. In adjusted analyses, differences in hypertension odds by insurance status were observed but were not statistically significant, suggesting that the relationship between insurance coverage and hypertension is likely influenced by additional factors. Effective chronic disease prevention and control require timely, high-quality, and longitudinal care that translates coverage into outcomes. Public insurance programs may play an important role in supporting prevention and management efforts by actively facilitating population health management for youth with emerging cardiometabolic risk. Strengthening care coordination, expanding community-based interventions, and reinforcing continuity of care across adolescence are essential next steps. Without additional preventive and care coordination strategies, increases in coverage alone may not translate into improved population health outcomes. This study's strengths include nationally representative, decade-long data with the latest NHANES cycle (2021–2023), clinically measured BP, and weighted analyses ensuring broad generalizability. Limitations include reliance on self-reported insurance data and a cross-sectional design that limits causal inference. Additionally, this study did not adjust for BMI as a key confounder that is known to influence hypertension risk and may contribute to observed trends.

Conclusion

The increasing burden of hypertension among U.S. youth highlights the critical role that health insurance, along with other socio-economic factors, play in facilitating timely access to preventive care. Medicaid and Medicare serve as essential safety nets, particularly for socioeconomically disadvantaged populations by providing comprehensive benefits such as blood pressure screenings, early interventions, and care coordination that would otherwise be inaccessible. As increasing youth obesity continues to drive early-onset hypertension and other metabolic disorders, the importance of sustaining and strengthening these programs remains high. Economic analyses demonstrate that Medicaid's investment in preventive care yields measurable returns, with children covered by Medicaid having better health outcomes and lower rates of hospitalization in adulthood, resulting in reduced long-term healthcare costs.[10-12]

Therefore, funding cuts to Medicaid, Medicare, or associated preventive health initiatives threaten to undermine decades of public health progress. Disruptions in coverage or access to affordable care during adolescence can widen existing health disparities and increase future healthcare expenditures. It is imperative that policymakers view Medicaid and Medicare as not just a mechanism for treatment affordability, but as foundational pillars for advancing early prevention, health equity, safeguarding youth

development, and reducing the nation's chronic disease burden over the life course.

References

1. Jackson SL, Zhang Z, Wiltz JL, Loustalot F, Ritchey MD, Goodman AB, Yang Q. Hypertension among youths—United States, 2001–2016. MMWR Morb Mortal Wkly Rep. 2018;67(27):758-762. Accessed July 15, 2026. https://www.cdc.gov/mmwr/volumes/67/wr/mm6727a2.htm
2. Charchar FJ, Prestes PR, Mills C, et al. Lifestyle management of hypertension: International Society of Hypertension position paper endorsed by the World Hypertension League and European Society of Hypertension. J Hypertens. 2024;42(1):23-49. doi:10.1097/HJH.0000000000003563
3. Chaturvedi A, Zhu A, Gadela NV, Prabhakaran D, Jafar TH. Social determinants of health and disparities in hypertension and cardiovascular diseases. Hypertension. 2024;81(3):387-399. doi:10.1161/HYPERTENSIONAHA.123.21354
4. National Center for Health Statistics. NHANES questionnaires, datasets, and related documentation. Centers for Disease Control and Prevention. Accessed July 15, 2026. https://wwwn.cdc.gov/nchs/nhanes/default.aspx
5. Centers for Disease Control and Prevention. NHANES survey methods and analytic guidelines. Accessed July 19, 2024. https://wwwn.cdc.gov/nchs/nhanes/analyticguidelines.aspx#sample-design
6. Di Bonito P, Di Sessa A. New diagnostic criteria for hypertension in children and adolescents: lights and shadows. Children (Basel). 2020;7(11):196. doi:10.3390/children7110196
7. Mendizábal B, Urbina EM, Becker R, et al. SHIP-AHOY (Study of High Blood Pressure in Pediatrics: Adult Hypertension Onset in Youth). Hypertension. 2018;72(3):625-631. doi:10.1161/HYPERTENSIONAHA.118.11434
8. Tsur AM, Akavian I, Derazne E, et al. Adolescent blood pressure and the risk for early kidney damage in young adulthood. Hypertension. 2022;79(5):974-983. doi:10.1161/HYPERTENSIONAHA.121.18748
9. Ostrovskaya MA, Rojas M, Kupferman JC, et al. Executive function and cerebrovascular reactivity in pediatric hypertension. J Child Neurol. 2015;30(5):543-546. doi:10.1177/0883073813494264
10. Wherry LR, Miller S, Kaestner R, Meyer BD. Childhood Medicaid coverage and later-life health care utilization. Rev Econ Stat. 2018;100(2):287-302. doi:10.1162/REST_a_00677
11. Brown DW, Kowalski AE, Lurie IZ. Long-term impacts of childhood Medicaid expansions on outcomes in adulthood. Rev Econ Stud. 2020;87(2):792-821. doi:10.1093/restud/rdz039
12. Maciosek MV, Coffield AB, Flottemesch TJ, Edwards NM, Solberg LI. Greater use of preventive services in U.S. health care could save lives at little or no cost. Health Aff (Millwood). 2010;29(9):1656-1660. doi:10.1377/hlthaff.2008.0701

REVIEW

Artificial Intelligence Applications in Orphan Disease Diagnoses and Drug Discovery

Sanidh Shah[1-2*], Ioannis Konstantinidis, PhD[3]

[1]Department of Biology & Biochemistry, University of Houston, Houston, Texas, USA
[2]Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia, Canada
[3]Department of Computer Science, University of Houston, Houston, TX, USA
[*]Corresponding author: Sanidh Shah; Email: sshah45@uh.edu
Disclosures: The authors report no financial relationships or conflicts of interest regarding the content herein.
Acknowledgements: We would like to thank Dr. Michelle Belco and Dr. Ioannis Konstantinidis, professors at the University of Houston, for their guidance and support in launching and supporting this project. Their advice and feedback were invaluable in the development of this review.

ABSTRACT

Background: Orphan diseases are diseases that are incredibly niche to a subset of the general population, affecting fewer than 1 in 2000 persons. Because they affect such a minor percentage of the population, they do not receive the attention or resources as those that plague the general population. Such diseases prove to be difficult to diagnose and potentially involve lengthy and cost intensive processes through multiple clinical trials. However, algorithmic advances offer a promising solution, attempting to leverage large scale probabilistic models to identify orphan diseases and/or generate novel orphan drugs.
Methods: This review evaluated recent research on the use of artificial intelligence (AI) for rare and orphan disease diagnoses and drug development, focusing on studies published between 2015–2025. A systematic search was conducted across four databases (PubMed, Scopus, Web of Science, Directory of Open Access Journals) following PRISMA 2020 guidelines, yielding seven studies for inclusion after screening.
Results: Studies demonstrate diverse AI applications for orphan diseases diagnoses and drug development spanning image-based, generative, and graph-/transformer-based architectures. Validation efforts in real-world biological or clinical settings remain limited with most studies relying on computational or in silico benchmarking. A lack of comprehensive, accessible background datasets was also apparent, with several applications serviced by proprietary or inaccessible data.
Conclusion: AI approaches show meaningful, albeit uneven, advances in orphan disease diagnostics across multiple algorithmic types, including image retrieval, generative data augmentation, and AI-assisted symptom checking. In comparison, drug discovery and repurposing applications show early but promising results but remain constrained without wet- lab or prospective clinical confirmation, highlighting a need for further research.

Keywords: Orphan diseases, Rare diseases, Orphan drugs, Drug development, Artificial intelligence

Introduction

Historical attempts surrounding orphan disease diagnoses and treatments have been constrained due to their rarity.[1] Diagnoses are achieved after lengthy consultation processes where patients are shuffled around between providers. After this "diagnostic odyssey," physicians may classify the disease as novel or existing based on scarce biomarker similarities identified. Orphan drug development follows this same timeline, with drug development processes proceeding thereafter. However, clinical trials are limited because of the reduced patient population size, leading to trials designed with weaker comparisons or with surrogate endpoints, compromising reliability. In supporting these efforts, the Orphan Drug Act of 1983 grants market exclusivity to firms that successfully develop a drug that passes the FDA's initial approval stages but even so, patients still face about 6.5 years for a successful orphan disease diagnosis and are often provided with expensive therapeutic options, if one is even available.[2,3] In this context, AI presents an opportunity to improve efficiency in orphan disease diagnoses and drug development/repurposing by enabling large-scale pattern matching.[4] Prior work has demonstrated AI's use in an eventual drug development and repurposing effort, drawing on both graph neural networks like TxGNN and transformer-based models like TxGemma.[5,6] However, because this use remains limited in scope in the diversity of past research, there remains a need to further understand its veracity and data bottlenecks.

Materials and Methods

This study was conducted following principles adapted from the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020 guidelines for systematic reviews. [7]

Research Question

The review addressed the following question: How are artificial intelligence models applied to rare and orphan disease diagnosis and drug development, and what empirical evidence exists regarding their clinical or biological validity given current data limitations?

Search Strategy & Database Description

A systematic literature search was conducted across electronic databases available through the University of Houston Libraries. Databases were selected to provide complementary coverage of biomedical, multidisciplinary, and specialized literature relevant to artificial intelligence applications in orphan diseases and orphan drugs.

- PubMed: Primary resource for biomedical literature and life science journals.
- Scopus: Multidisciplinary database providing broad coverage of health sciences, engineering, and related disciplines.
- Web of Science: Citation database included to supplement retrieval and reduce database-specific indexing bias.
- Directory of Open Access Journals: Open-access journal index included to improve identification of niche literature relevant to orphan drugs not comprehensively represented in traditional databases.

The final search was performed in December 2025 and included publications from January 2015 to December 2025. Two search approaches were taken for orphan/rare disease diagnosis and orphan drug development:

- Approach 1 (Diagnosis): Targeted for primary research using Boolean string below.
- Approach 2 (Drug Discovery): Relaxed eligibility to include secondary literature and conceptual narratives due to the current scarcity of primary data in the subdomain.

Search queries were constructed using the following Boolean operators:

- ("artificial intelligence")
  AND
- ("rare diseases" OR "orphan diseases")
  AND
- ("diagnosis" OR "drug discovery" OR "drug repurposing")

Inclusion and Exclusion Criteria

In this review, rare diseases are identical in meaning to orphan diseases and studies not properly limited to such diseases were excluded. Studies were included if they met all the following criteria:

- Peer-reviewed research articles in English
- Published between 2015–2025
- Explicit application of AI models to rare/orphan disease diagnosis or orphan drug repurposing/use
- Reported empirical results (model performance, validation, or clinical relevance)

Studies were excluded if they:

- Were editorials, commentaries, or opinion pieces

unless they addressed drug development where no primary data existed.
- Focused solely on common diseases.
- Did not specify datasets or model architectures.
- Were purely theoretical with no implementation.

Study Selection & Review Process
A single reviewer conducted all stages of the review process, including title/abstract screening and full-text eligibility assessment. Interrater reliability (IRR) measures, like Cohen's Kappa, cannot therefore be calculated and the risk of individual bias in study selection and data extraction is present as a limitation of this paper. Each included study was selected using the adapted framework in this section and results were recorded in a standardized assessment matrix for uniform application of criteria across all seven studies.

Data Extraction
From each included study, the following variables were analyzed:

- Disease domain
- AI model type (e.g. CNN, GAN, GNN, diffusion, transformer)
- Data set size and source
- Data modality (imaging, tabular, genomic, EHR, synthetic)
- Validation method (internal, external, retrospective, none)
- Reported performance metrics (accuracy, sensitivity, AUC, etc.)
- Evidence of real-world or clinical validation

Quality and Bias Assessment
Given the heterogeneity of study designs, methodological quality was assessed using an adapted framework based on:

- Dataset transparency and accessibility
- Sample size adequacy
- Validation rigor
- Risk of data leakage or synthetic bias
- Presence of clinical or biological ground truth

For primary research, all five criteria were applied. For conceptual or narrative reports in the drug discovery tier, the assessment focused on the biological validity of the proposed AI framework and the quality of evidence supporting the drug-gene linkages. Studies were qualitatively classified as low, moderate, or high methodological reliability.

Synthesis Strategy
Due to high heterogeneity in AI architectures, outcome measures, disease types, and validation methodologies, a formal statistical meta-analysis was not feasible. Instead, a qualitative meta-synthesis was chosen to identify architectural trends and data bottlenecks that quantitative metrics alone might obscure, grouping studies into:

- AI-assisted diagnosis
- Synthetic data generation
- Drug repurposing via GNNs and transformer-based platforms

Results

Study Selection
Database searching yielded 202 records after duplicate removal. Following title/abstract screening and full-text eligibility assessment, seven studies were included in the final review. Reasons for exclusions are detailed in the flow diagram (Figure 1).

Study Characteristics
The final review consists of seven studies across a wide setting, including Italy, Germany (University Hospital Essen), Canada (University of Toronto), and the United States (Harvard University) and discusses a variety of rare/orphan diseases including the Fabry disease, hemophilia, hereditary cerebellar ataxia, and rare pulmonary diseases.

Usage of Artificial Intelligence in Orphan Disease Diagnosis
To improve clinical outcomes, artificial intelligence has been proposed as complementing physician diagnoses through their setup as so-called "symptom checkers." Symptom checkers are LLMs that provide likely diagnoses based on a patient's inputting their symptoms. However, while patients report their interactions with such LLMs as generally positive and assistive to their care, physicians hold a more tempered reaction with 84% of general practitioners (GPs) in an Italian study reporting the AI as insignificant to patient care.[8]

Early Diagnostic Assistance in Anderson-Fabry Disease. Addressing these reactions, a mixed-methods study aimed to determine whether integrating expert interview vignettes into AI-powered symptom checkers could enhance diagnostic accuracy for rare diseases, with Fabry disease as the primary focus.[9] Artificial intelligence was demonstrated as capable of providing first-time right checks by implementing expert vignettes onto model training in complementation to

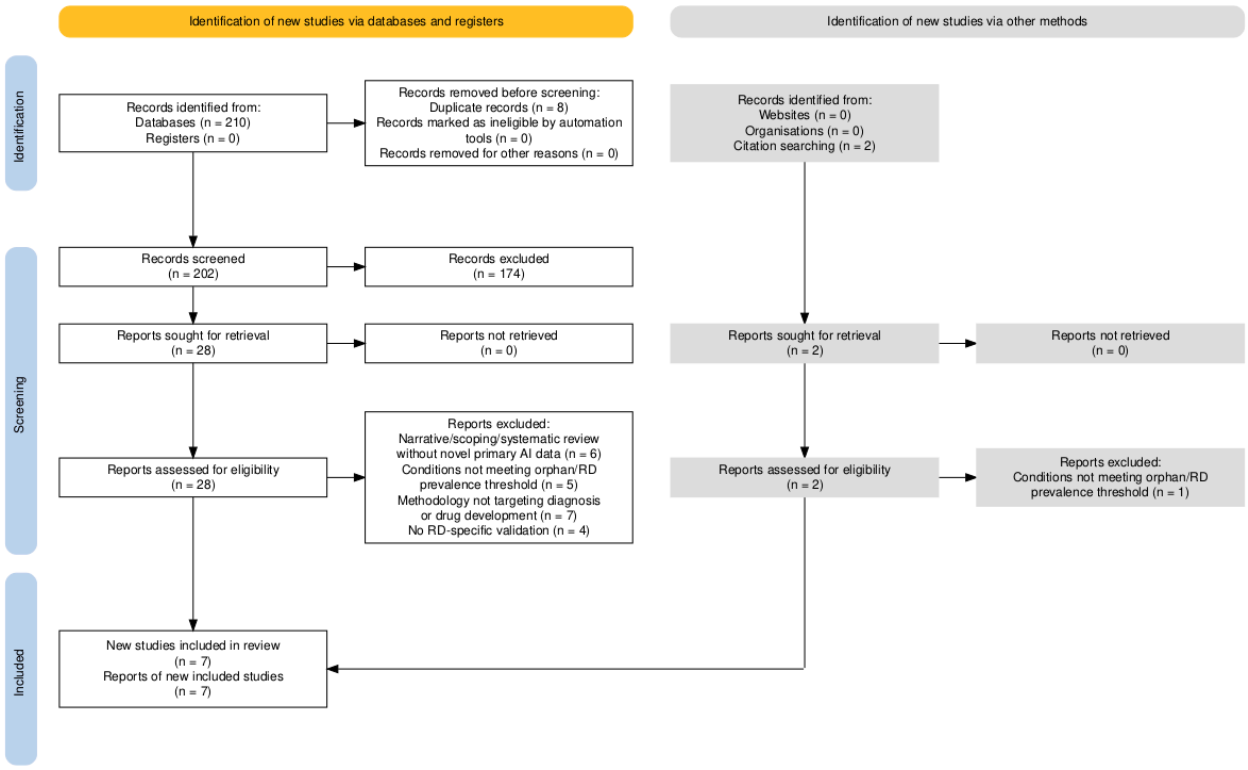

Figure 1. PRISMA flow diagram of study selection.[10]

physician diagnosis. By supplying these clinical vignettes, the authors successfully improved the AI model's correct inferences on the Fabry Disease by 16% compared to the non-vignette AI model. However, the overall efficacy of AI remained limited, with a 33% accuracy rate despite overall improved patient satisfaction remarks, suggesting that AI-assisted diagnostic procedures promoted patient satisfaction at the risk of contradicting their primary care provider. These results were derived from a proprietary model, Ada DX, an AI-powered symptom checker developed by Ada Health GmbH (Berlin, Germany) representing limited dataset accessibility.

Use of AI-Assisted Imaging. Analyzing medical imaging outputs can lead to significant variation in physician interpretation and may produce errors in eventual diagnostic processes.[11] Addressing these concerns, artificial intelligence has been successfully used in studies making use of medical imaging techniques, including computerized tomography (CT) and ultrasound. In one study conducted at University Hospital Essen in Essen, Germany, artificial intelligence was used to supplement resident physician diagnostic CT readings, aiming to determine whether residents in radiology could provide more accurate diagnoses of chest CTs using content-based image retrieval (CBIR) searches without substantial increases in diagnosis time.[12] Using a 150-point grading system, researchers evaluated clinicians' diagnostic accuracy with and without the AI search engine, yielding scores of 17.6 points unassisted and 32.0 points assisted. Concurrently, the time to diagnosis was also significantly reduced with the implementation of the tool, with compounding effects as physicians became further familiarized. However, several things were of concern: 1) The need for a large database, 2) Lack of standardized control group, and 3) Slice thickness limitation. In developing the search engine, the authors discuss the generation of the image database from only

621 patients, highlighting the need for sourcing sufficiently large amounts of data for further evaluations of algorithmic models independently undertaken.

Nonetheless, by using CBIR processes, AI is shown to complement physician diagnoses of rare diseases with image matching and validation checks, as well as affording logistical benefits through time savings. In acquiring their data, the authors' jointly developed a CBIR similar patient search (SPS) system with Siemens Healthineers (Erlangen, Germany) utilized a reference database of CT scans from 621 patients afflicted with 1/69 possible pulmonary diseases. For each CT scan loaded, target regions of interests (ROI) for a particular disease were marked by an expert chest radiologist with a total of the abovementioned 13,658 ROIs included in the database.

Generation of Synthetic Imaging for Hemophilia. In contrast to content-based image retrieval (CBIR) systems, generative models also demonstrate potential in synthetic medical imaging, as explored in a generative ultrasound imaging study for 12 hemarthrosis in hemophilia patients in processes that would otherwise need thousands of data points from non-existent patients.[13] AI models can generate broad understandings of the rare disorder, termed domain knowledge, subsequently leveraged to synthesize further additional data points in the form of ultrasound images. These synthetic ultrasound images are then used to train a downstream convolutional neural network (CNN) classifier for hemarthrosis detection, resulting in improved diagnostic sensitivity and overall classification performance of the CNN compared with training on real data alone, particularly at low sample sizes. However, although visually comparable, synthetic ultrasound images achieved lower CNN classification accuracy than real images, with 75.1% versus 84.9% agreement with ground-truth annotations.

Additionally, neural models trained on the synthetic images failed to consistently segregate between control images and the target hemarthrosis, with intermittent misclassification of control recess distension images as hemarthrosis-present condition. Even so, the synthetic generation framework was able to achieve an overall real-like image quality, with a substantial improvement in diagnosis performance marked by a 27% sensitivity increase compared to other downsampling or typical augmentation.

The study utilized a content generator, Gc, to create diverse ultrasound images from the base dataset, Db, which represented 11,676 images sourced of 3,710 knee examinations of non- hemophiliac patients, modified by a context generator, Gs, to adapt these first-run synthetic images to target dataset, Dt, which consisted of 360 images from 67 knee examinations from hemophiliac patients.

Generation of Synthetic Datapoints through Generative Adversarial Networks for pwCA. Generative artificial intelligence may also be used to augment small sample groups and "balance" unbalanced datasets resulting from insufficient datapoints for specific rare diseases. While traditional balancing methods–including undersampling, oversampling, and Synthetic Minority Over-sampling Technique (SMOTE)–can successfully generate additional data points to improve classifier training, they remain limited in accurately modeling the complex, high- dimensional probability distributions of the original rare disease data. To better model these distributions, effects of advanced data augmentation techniques, such as generative adversarial networks (GANs), were investigated on gait datasets associated with primary hereditary cerebellar ataxia (pwCA).[14] Consequently, generative AI outputs, specifically conditional Tabular Generative Adversarial Networks (ctGANs), were found to be most effective at enabling a classifier to correctly identify true positive cases that were underrepresented in the initial unbalanced dataset. The researchers procured two samples of 30 pwCA patients and 100 age and gait matched healthy subjects from Latina, Italy over a period of roughly 1.5 years. Specific spatiotemporal and kinematic gait characteristics were measured using inertial measurement units (IMU) sensors and extracted using G-STUDIO, with further parametric extractions, data cleaning, preprocessing, and augmentation. Generative AI methods were then compared to traditional methods in an evaluation of five data balancing strategies, comprising two AI and three traditional non-AI approaches. A random forest classifier was subsequently trained on the augmented datasets to distinguish between pwCA and non-pwCA gait data. Implementations of GANs and ctGANs followed, with only ctGANs outperforming all other methods, achieving 90% accuracy for N = 200 and 81% accuracy for N = 1000, with minimal log losses (0.35 and 0.40, respectively). Conversely, standard GANs did not significantly improve performance over traditional methods. While the exclusive focus on gait parameters excludes other clinical features and limits short-term validation, the study demonstrates the potential of conditional generative modeling for tabular biomedical data, although such approaches may not generalize to

other data types, such as medical images, which require spatially aware generative architectures, as explored in prior rare disease data synthesis work.[13] The authors note that data were independently sourced in Italy and may be available upon request.

Drug Development & Repurposing for Orphan Diseases

Developmental Barriers. While efforts to treat orphan diseases have been spurred through legislative efforts, development of drugs for such illnesses have generally remained limited due to insufficient financial viability, limited clinical data, and high levels of complexity.[15] Here, AI is purported to deliver cost reduction by "streamlining and improving efficiency," improving cost margins and expenses as opposed to reducing financial risks.[16] Such algorithms can generate likely outcomes of clinical trials for a given orphan drug, acting as a data surveyor for appropriate matches. Patient data acquisition processes are accelerated, tracking down suitable patients from electronic health records (EHRs). Contextually, in a separate study, 413 patients were identified as suitable match for using a clinical trial for a rare disease phenotype using a computational phenotype algorithm compared to 179 patients being identified via a traditional registry, although in contrast to LLM-inspired models largely explored here.[17]

AI is also shown to address the high level of overall complexity in developing orphan drugs by assisting in the drug research process, with a key finding in the drug-gene link. Large scale biomedical repositories, such as the Connectivity Map (CMAP) can be serviced by LLM models, training them to search for orphan drugs via a drug-gene relationship.[16] This study however remains limited with few research articles cited in the current utilization of artificial intelligence and with significant levels of assumptions made in general commentary but nonetheless serves as a rendition of how AI may be used in the direct drug production pathway.

AI-Augmented Clinical Trials. AI also poses hypothetical applications in joint diagnoses and drug repurposing in consumer- grade setups.[18] The use of a LLM, in chatbot form, is imagined to "predict" what orphan disease the patient might be afflicted with based on symptom data, subsequently linking to a clinical decision support system (CDSS), suggesting possible therapeutics based on an amalgamation of the internal disease mode, case registry, research data repository, and hospital health data records. While such theoretical applications do not discuss the sourcing of a dataset, the published narrative does confirm specific ideations surrounding the ability to use AI models to utilize pre-established datasets to 1) suggest disease match and 2) propose possible therapeutic agents, the latter which is of a greater interest for this section.

Graph Neural Network Assisted Drug Repurposing. Some efforts have successfully bridged the use of AI in repurposing therapeutic agents, most notably with the introduction of TxGNN, a graph foundation model targeted to rare/orphan diseases. In contrast to personalized current drug and patient information, TxGNN operates on heterogeneous medical knowledge graphs to rank drugs as "potential indications and contraindications" alongside a metric learning module, tracing relationship matches on molecular interactions, chemical similarities, effect associations, phenotypic results, and overall pathway involvements.[6] This model uses a dataset of 17,080 diseases as its reference library, with nearly 8,000 potential therapeutics in development as a supplement, comprising the TxGNN Predictor module. As suggested by its name, the predictor module is responsible for predicting the drug disease relationship, leveraging its reference database to score pairings:

$$s\ (d, x, r)$$

These pairings reconcile a given drug (d) and disease (x) based on the type of relation–whether a contraindication or an indication (r)–with higher scores meaning a better match and therefore a likely therapeutic candidate and vice versa with lower scores. To ensure translatability of drug candidates from treated diseases to non-treated diseases, a metric learning module was also integrated, fitting within the predictor training via counterfactual questions, to enforce disease/indication and disease/contraindication distances and thereby better shape the geometry of the predictor's resulting embeddings. TxGNN's Explainer module follows these mathematical relations by exposing multi-hop qualitative evidence linking the drug to disease, suitable for human reasoning and analysis of the model's proposed linkages.

Clinical validity was tested by considering three orphan diseases: Kleefstra's syndrome, the Ehlers-Danlos syndrome, and Nephrogenic Syndrome of Inappropriate Antidiuresis (NSIAD). TxGNN successfully predicted repurposed drug candidates for all three diseases, all verified as suitable candidates by supporting retrospective validation via electronic medical records (EMR). However, candidate drugs

were cross evaluated against physician prescriptions and observations, introducing confounding as predictions may reflect learned prescribing patterns rather than true therapeutic efficacy. Causality thus cannot be proven, and validation attempts remain primarily observational. The medical knowledge database is additionally identified as a greatly limiting factor, which could possibly include data biases and outdated references, and this approach remains fundamentally different from the previously explored AI approaches employing patient data or literature mining at inference time. This database was sourced from a variety of biomedical databases, including DrugBank, SIDER, DisGeNET, KEGG, etc. and indicate that it is available on Harvard University's Dataverse tool under a persistent identifier.

Transformer Assisted Drug Synthesis. LLM advancements introduce another viable domain for identifying and developing therapeutics for orphan diseases compared to graph-based AI models. Although outside the scope of the systematic synthesis, Google DeepMind's TxGemma demonstrates how transformer algorithms may support multimodal therapeutic reasoning, molecular prediction, and agentic drug-discovery workflows, which may be applicable to orphan diseases from its current common disease profile.[5] However, validation efforts here similarly remain primarily in silico with limited independent experimental validation reported to date, especially with regard to orphan diseases.

Discussion

Studies focusing on the use and application of artificial intelligence regarding orphan diseases appear to center around diagnoses, most principally through synthetic data creation, symptom checkers, image/numerical detection, and possibly prognosis extrapolation. These efforts cluster out in varying methodologies but demonstrate the expansive capability of AI systems. Generally, various AI models can deliver with reasonable accuracy and reliability, correct and successful diagnoses.

However, dataset availability remains a challenge in verifying these capabilities, especially in the instance of synthetic data creation. Several studies were conducted using private, proprietary models, representing limited data accessibility. Public institutional studies making use of independently sourced data generate some leads to data acquisition, paving the way for future studies. Regarding data sourced by commercial enterprises, such firms provided access through integrated AI platforms as in the case of Ada Health GmbH and Siemens AG. However, two independently-ran studies discuss their dataset availability as being available on request to the authors.[12,14] Additional studies also disclose the use of independent datasets, without indication to their accessibility.[13] While such acquired datasets may then be limited in comparison to those sourced by larger commercial firms, they represent a potentially high level of accessibility for future studies.

In the development of orphan drugs, these dataset challenges also pose a difficulty in generating viable drug candidates, both in repurposing and de novo synthesis. As demonstrated in the case of TxGNN, limited background data can result in confounding and data biases. Transformer AI models offer a promising solution; however, there exists no wet lab validation of such models in generating therapeutic matches for diagnosed orphan diseases. Current literature purports that genomic data, arranged in graphical views, are best suited for graph neural networks whereas diverse source data is best served by transformer models, which include LLMs and are representative of the latest advances in the field of AI. Hence, new advances are likely best served by transformer models.

In combating these dataset limitations, and beyond the formal search process, larger biomedical repositories like Orphanet, Cyagen, and the UK Biobank may serve as accessible entry points.[19–21] These aggregates link to individual collections of datasets, organized by country of origin and organization. Some existing projects such as RareSim (Zürich, CH) implement the use of such datasets in AI-assisted setups, specifically the UK Biobank, leveraged with a similarity AI model to track patient cases to build a more comprehensive disease profile.[22] However, AI-assisted workflows for therapeutic development using these aggregates have not yet been demonstrated in the literature.

Future Work

Because significant work exists in using both AI and traditional research pathways to identify and diagnose orphan diseases, it is beneficial to analyze orphan drug development and drug repurposing. Building on the lack of sufficient wet-lab validation in drug repurposing demonstrated across reviewed studies, particularly in case of TxGNN's retrospective chart review, future work may focus on both biological and clinical evaluation of AI recommended therapeutics, especially those suggested by transformer models.[5,6] LLMs have been greatly introduced into the public domain, popularized by ChatGPT, Gemini, and Claude. Hence, there is a future prospect of the use of such

models to predict drug candidates in the near future in real-world settings.[16] Dataset availability challenges highlight potential concerns on the validity of these inferences and promote additional research in benchmarking.[15]

Therefore, based on observed gaps in dataset accessibility and the limited validation of AI to orphan drug repurposing, a critical question arises: "How do graph-based and transformer-based AI models compare in reliably mapping orphan-disease genomic and symptom profiles to appropriate existing therapeutics, and how closely do these recommendations align with expert clinical practice?"

To investigate this, one selected GNN model and one transformer-based model may be chosen for a comparative study. To ensure a "zero-shot" evaluation, AI models may be blinded to specific data linking a treated orphan disease to its known therapy. Models may then be instructed to rank commercially available drugs for an inputted orphan disease based on associated symptomatic and genomic data. The methodology may focus on directing structural, graph-based data to GNNs and wide-ranging information deposits to transformer models to leverage their respective architectural strengths consistent with current recommendations in the field.[5,6,16]

The resulting ranked-choice outputs may then be compared to traditional physician treatment courses, verified by clinical trials, serving as efficacy benchmarks. Weighted similarity scores (concordance) may then be used to evaluate the models' outputs against expert selections, scoring on metrics of direct match (concordance rates), chemical composition similarity, and functional match (drug-target binding affinity) building on validation frameworks proposed in prior orphan drug repurposing efforts.[6,16]

Conclusion

Artificial intelligence has been successfully implemented in both the diagnosis of orphan diseases and the development and repurposing of existing therapeutics as orphan drugs. For diagnoses, reviewed work demonstrates the capability of diverse AI architectures– including CNN-based image retrieval, generative adversarial networks, and transformer-based symptom checkers– across symptom pattern analysis (SPS), image processing, and data inference. For drug development and repurposing, reviewed literature spans computational phenotyping for improved patient identification, conceptual LLM-CDSS frameworks for therapeutic suggestion, and graph neural network-based repurposing via TxGNN, with transformer-based models such as TxGemma representing a promising and emerging direction outside the scope of the systematic synthesis.[5,6,17,18]

In all cases, dataset availability was generally limited, with several diagnostic studies being privately sourced, institutionally prepared, or undisclosed. In the case of GNN models, the vulnerability of graph-based models to biases and confounding is expressly highlighted. In contrast, transformer based models, for both orphan disease diagnoses and drug development/repurposing, were shown to be more resistant to limited dataset challenges, with distinct data synthesis capabilities.

However, AI assisted drug development and repurposing have been limited with only in silico validation and retrospective chart review in the case of TxGNN. No wet lab validation has yet been conducted.

References

1. Jefferies N. Diagnostic odyssey in rare disease. Knowledge Hub. Published online. Accessed October 29, 2025. https://www.genomicseducation.hee.nhs.uk/genotes/knowledge-hub/the-diagnostic-odyssey-in-rare-disease/
2. Mikami K. Orphans in the market: the history of orphan drug policy. Soc Hist Med. 2019;32(3):609-630. doi:10.1093/shm/hkx098
3. Roberts AD, Wadhwa R. Orphan Drug Approval Laws. In: StatPearls. StatPearls Publishing; 2023.
4. Xu N, Yang D, Arikawa K, Bai C. Application of artificial intelligence in modern medicine. Clin eHealth. 2023;6:130-137. doi:10.1016/j.ceh.2023.09.001
5. Wang E, Schmidgall S, Jaeger PF, et al. TxGemma: efficient and agentic LLMs for therapeutics. arXiv. Published online April 2025. doi:10.48550/arXiv.2504.06196
6. Huang K, Chandak P, Wang Q, et al. A foundation model for clinician-centered drug repurposing. Nat Med. 2024;30(12):3601-3613. doi:10.1038/s41591-024-03233-x
7. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi:10.1136/bmj.n71
8. Mahlknecht A, Engl A, Piccoliori G, Wiedermann CJ. Supporting primary care through symptom checking artificial intelligence: a study of patient and physician attitudes in Italian general practice. BMC Prim Care. 2023;24:245. doi:10.1186/s12875-023-02143-0
9. Pankow A, Meißner-Bendzko N, Kaufeld J, et al. Medical expert knowledge meets AI to enhance symptom checker performance for rare disease identification in Fabry disease: mixed methods study. JMIR AI. 2025;4:e55001. doi:10.2196/55001
10. Haddaway NR, Page MJ, Pritchard CC, McGuinness LA. PRISMA 2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. Campbell Syst Rev. 2022;18(2):e1230. doi:10.1002/cl2.1230
11. Quinn L, Tryposkiadis K, Deeks JJ, et al. Interobserver variability studies in diagnostic imaging: a methodological systematic review. Br J Radiol. 2023;96(1147):20220972. doi:10.1259/bjr.20220972
12. Haubold J, Zeng K, Farhand S, et al. AI co-pilot: content-based image retrieval for the reading of rare diseases in chest CT. Sci Rep. 2023;13(1):4336. doi:10.1038/s41598-023-29949-3

13. Mendez M, Castillo F, Probyn L, Kras S, Tyrrell PN. Leveraging domain knowledge for synthetic ultrasound image generation: a novel approach to rare disease AI detection. Int J Comput Assist Radiol Surg. 2025;20(2):415-431. doi:10.1007/s11548-024-03309-6
14. Trabassi D, Castiglia SF, Bini F, et al. Optimizing rare disease gait classification through data balancing and generative AI: insights from hereditary cerebellar ataxia. Sensors (Basel). 2024;24(11):3613. doi:10.3390/s24113613
15. Belousova OA, Groen AJ, Ouendag AM. Opportunities and barriers for innovation and entrepreneurship in orphan drug development. Technol Forecast Soc Change. 2020;161:120333. doi:10.1016/j.techfore.2020.120333
16. Irissarry C, Burger-Helmchen T. Using artificial intelligence to advance the research and development of orphan drugs. Businesses. 2024;4(3):453-472. doi:10.3390/businesses4030028
17. Geva A, Gronsbell JL, Cai T, et al. A computable phenotype improves cohort ascertainment in a pediatric pulmonary hypertension registry. J Pediatr. 2017;188:224-231.e5. doi:10.1016/j.jpeds.2017.05.037
18. Ronicke S, Hirsch MC, Türk E, Larionov K, Tientcheu D, Wagner AD. Can a decision support system accelerate rare disease diagnosis? Evaluating the potential impact of Ada DX in a retrospective study. Orphanet J Rare Dis. 2019;14(1):69. doi:10.1186/s13023-019-1040-6
19. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3):e1001779. doi:10.1371/journal.pmed.1001779
20. Singh J. The portal for rare diseases and orphan drugs. J Pharmacol Pharmacother. 2013;4(2):168-169. doi:10.1177/0976500X20130205
21. Rare Disease Data Center (RDDC). About RDDC. Research Institute of Tsinghua, Pearl River Delta; Cyagen Biosciences. Accessed May 18, 2026. https://rddc.tsinghua-gd.org/about
22. UK Biobank. RareSim: similarity-based AI for diagnosing rare diseases. Accessed May 18, 2026. https://www.ukbiobank.ac.uk/projects/raresim-similarity-based-ai-for-diagnosing-rare-diseases/

REVIEW

Bridging the Gap: A Narrative Review of Surgical and Specialty-Specific Bootcamps in Easing the Transition to Residency

Major Eldridge, BS[1*], Amelia Tower, DO[1]

[1]Naresh K. Vashisht College of Medicine, Texas A&M University, Bryan, TX, USA
[*]Corresponding author: Major Eldridge; Email: majoreldridge@tamu.edu
Disclosures: The authors report no financial relationships or conflicts of interest regarding the content herein.

ABSTRACT

Background: The objective of this study is to evaluate the role of surgical and specialty-specific bootcamps in improving preparedness for the transition from final-year medical student (MS4) to postgraduate year 1 (PGY-1).

Methods: The study design is a narrative review of studies published from 2010-2025 identified through PubMed and Google Scholar. Articles were included if they described surgical and specialty-specific bootcamps involving MS4 or PGY-1 trainees and reported outcomes on procedural competence, confidence, or clinical knowledge. A final total of 12 articles met inclusion criteria and were included in this review. The study setting spans medical schools and residency programs in the United States, and participants represent studies that include medical students, sub-interns, and incoming residents participating in surgical or specialty-specific bootcamps.

Results: Medical school-based bootcamps improved procedural competency and confidence. Sub-internship bootcamps enhanced preparedness and reinforced specialty-specific milestones. Residency-level bootcamps facilitated rapid skill acquisition, with longitudinal practice improving retention. Specialty-specific bootcamps, including ophthalmology and dermatology, improved objective clinical performance and confidence.

Conclusion: Bootcamps at every stage of training significantly strengthened immediate clinical readiness and technical competence. Most studies involved small, single-institution cohorts and focused on short-term outcomes. Despite the limitations of individual studies, the consistency of findings across multiple institutions and specialties strengthens the overall credibility of the evidence. Broader adoption of bootcamps, combined with long-term reinforcement, could mitigate the steep learning curve at the start of general and subspecialty surgical residency. Future research should assess whether the early gains in skill and confidence reported in these programs persist across institutions and translate into measurable improvements in residency performance and patient care.

Keywords: Transition to residency, Surgery, Bootcamp, Specialty, Medical education

Introduction

Medical students and residents strive to be well-prepared for the next stage of their training.[1] Similarly, residency program directors seek team members who are coachable learners with a solid baseline of skills and knowledge.[2] Modern hospital systems prioritize efficiency, patient safety, and quality outcomes.[3] The transition from final-year medical student (MS4) to postgraduate year 1 (PGY-1) represents a significant gap in medical education, with implications for all stakeholders.

Despite the stress of away rotations, ERAS applications, and The Match ®, MS4s often experience a less structured final year of undergraduate medical education (UME). While this period offers important mental and physical reprieve, students also seek preparation for the demanding years of residency, particularly in surgical fields. Likewise, the American College of Surgeons (ACS), American Board of Surgery, and Association of Programs Directors in Surgery collectively released a call-to-action statement highlighting the need for UME to provide additional preparation for surgical residency.[1] To bridge this gap, medical schools and residency programs have implemented surgical bootcamps designed to enhance procedural skills, clinical confidence, and readiness for residency. For the purposes of this review, bootcamps are defined as short, intensive educational interventions designed for students to rapidly gain procedural skills, clinical understanding, and confidence in preparation for a new stage in training.

This literature review evaluates the common practices, effectiveness, and limitations of 12 established general and subspecialty surgical bootcamps in the United States (U.S.). These programs are hosted at one of three sites: the MS4 home medical school, a subinternship site, or the PGY-1 residency program. While this review focuses on U.S. based bootcamps, similar initiatives exist internationally in Canada, the United Kingdom, and France. Differences in medical education structures limit direct comparison, but their existence underscores the global recognition of the challenges associated with surgical residency preparedness.[4,5,6]

Materials and Methods

A narrative review was conducted to identify published studies related to the efficacy of surgical skills bootcamps in easing the transition to residency. This methodology was chosen to characterize the breadth of literature on surgical bootcamps across the stages of medical education and to identify patterns, gaps, and differences in implementation.

Database searches were performed from July 2025 to August 2025 using PubMed as the primary database. Google Scholar was used as a supplementary source to identify additional relevant studies and to retrieve full-text articles not indexed in PubMed, such as Joshi et al. Search terms included “transition to residency,” “surgical,” “bootcamp,” “subspecialty,” “specialty-specific,” and “residency readiness,” used in relevant combinations with Boolean operators (AND, OR).

Titles, abstracts, and full texts were screened by a single reviewer using predefined inclusion and exclusion criteria. Articles published between 2010 and 2025 were included if they reported on general and specialty-specific surgical skills bootcamps within the U.S. at the medical school, sub-internship, or residency level and described outcomes on procedural competence, confidence, or knowledge. Only studies involving MS4, subintern, or PGY-1 populations were included. Commentaries, editorials, and non-peer-reviewed articles were excluded.

The initial search yielded 13 studies, of which 9 met inclusion criteria after full-text review. Three studies were excluded for being conducted outside the U.S. and one was excluded for not meeting the definition of a bootcamp. Subsequent cross-referencing of prior reviews identified 3 additional articles that met inclusion criteria, resulting in a final total of 12 included studies. The study screening and inclusion process is illustrated in Figure 1.

Results

This review synthesizes published outcomes of surgical and specialty-specific bootcamps across three educational stages: MS4, subinternship, and PGY-1. A summary of study characteristics and primary findings is provided in Table 1.

MS4 Home Medical School Based Bootcamp

Medical schools hosting surgical bootcamps provide a final opportunity to enhance graduating students’ readiness for residency. Most programs focus on general surgery knowledge and technical skills rather than specialty-specific training. Outcomes typically include both objective measures of procedural competency and self-reported confidence and preparedness.

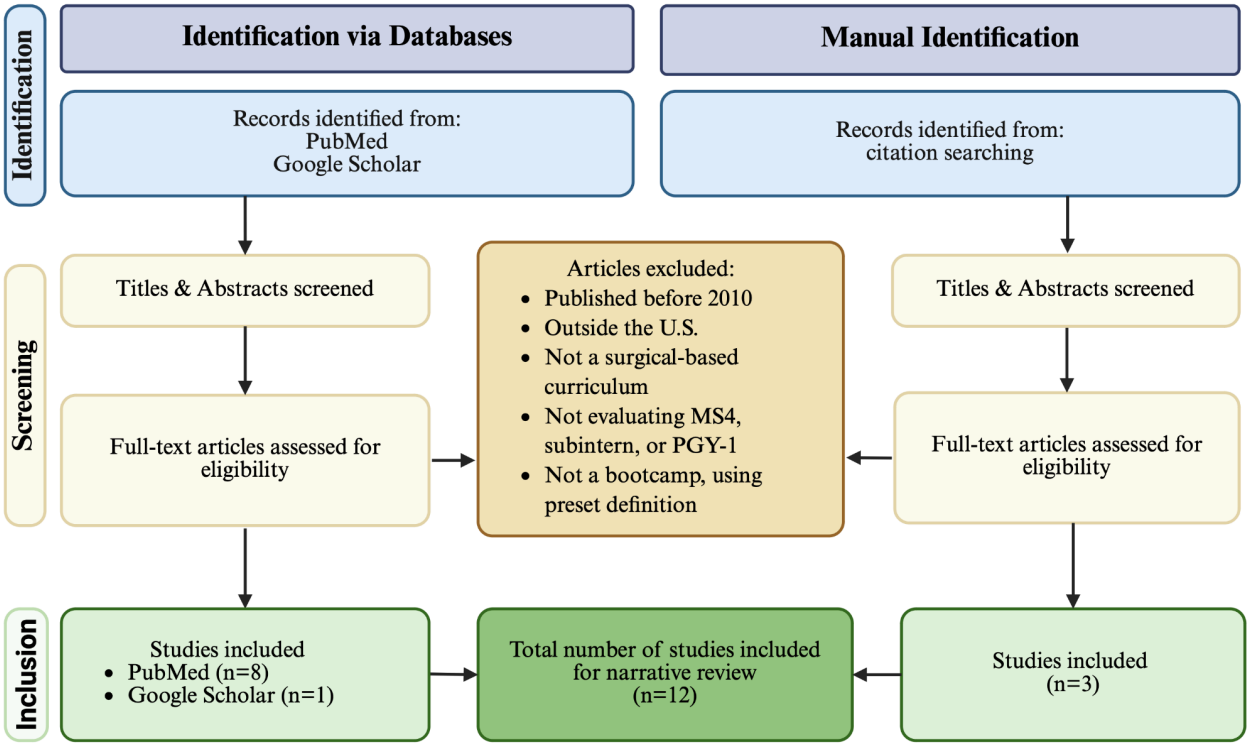

Figure 1. Illustration of article selection and exclusion methodology.

The University of Pennsylvania School of Medicine introduced a pregraduation surgical bootcamp to assess whether MS4 confidence and technical skills could improve. Twenty-eight students, entering general surgery or a surgical subspecialty, participated in the five-day course and completed anonymous surveys before, immediately after, and six months post-bootcamp. The course covered clinical scenarios and physical skills such as, chest tube insertion, cricothyroidotomy, and nasogastric (NG) tube placement. A control group of nine surgical interns who did not participate received the same survey six months into surgical internship. Immediately following the bootcamp, participants reported greater confidence in all skills except peripheral intravenous catheter placement. Six months into residency, the bootcamp cohort demonstrated higher confidence than controls in performing a cricothyroidotomy ($p = 0.04$) and inserting a chest tube ($p = 0.05$).[7]

Stony Brook University and Thomas Jefferson University Hospitals implemented a two-week surgical bootcamp for MS4s matching into surgical subspecialties. Students were scored on five tasks consisting of patient handoff, suturing, knot tying, central line placement, and chest tube placement using objective rating tools from the ACS before and after the bootcamp. Two confidence measures were also collected pre- and post-course. Among the 12 participants, objective skills scores and self-reported confidence improved in all five tasks ($p < 0.01$).[2]

The University of Miami Miller School of Medicine hosts a four-week surgical bootcamp. The MS4s attend didactic lectures on common surgical consultation scenarios and perioperative care, alongside practical workshops designed by resident and attending educators. Skills taught include suturing, knot tying, laparoscopic and robotic surgical procedures, and bedside/ICU procedures. Workshops are delivered one-on-one, fostering individualized assessment and feedback. Over two years, 28 participants completed

the bootcamps and associated pre and post-bootcamp surveys. Self-reported confidence and residency preparedness improved following the bootcamp (P < .001).[1]

The University of Michigan Medical School implemented an innovative five-day surgical skills bootcamp for MS4s matching into surgical residencies. Under the guidance of faculty surgeons, 23 students practiced abdominal, thoracic, and vascular surgical procedures on cadavers. Evaluation by faculty and the learners themselves showed increases in procedural knowledge and confidence ($p < 0.05$). Additionally, comparison of pre- and post-bootcamp data showed decreases in self-reported anxiety ($p < 0.05$).[8]

Collectively, these studies demonstrate the benefits of medical school-hosted surgical bootcamps. Confidence and technical skill ability improved markedly, and students reported enhanced preparedness for residency. Most studies were limited by small sample sizes and single-institution design.

Subinternship Based Bootcamp

The subinternship, or acting internship, often serves as the bridge between MS4 and PGY-1. This rotation allows students to perform medicine at a level akin to that of an intern. For students focused on surgery, subinternships are highly formative and often essential for those pursuing competitive surgical subspecialties. This experience offers opportunities for didactic instruction, interprofessional leadership, and specialty exploration that extend beyond the standard core surgery clerkship.[9] These features create an ideal environment for embedding a surgical bootcamp.

The Ohio State University Medical Center Department of Urology hosted 29 students for a subinternship-based bootcamp. Over the four-week rotation, students engaged with an internally developed online module curriculum and participated in a half-day simulation workshop. The curriculum aligned with Accreditation Council for Graduate Medical Education (ACGME) Level 1 milestones for urology. The simulation session included tasks such as 16Fr Foley catheter placement, bladder irrigation, diagnostic cystoscopy, and advanced Foley catheter placement over wire. The authors defined "objective knowledge" as performance on medical and procedural knowledge assessments administered before and after the online modules, while "subjective knowledge" was self-reported confidence and skill mastery measured on a five-point Likert scale. Subinterns demonstrated gains in both objective and subjective knowledge following the curriculum ($p < 0.001$, respectively). Participants reported the bootcamp enhanced their understanding of urological topics ($p < 0.001$), prepared them to meet ACGME milestones ($p < 0.001$), and would recommend the curriculum to future subinterns ($p < 0.001$).[10]

The included study demonstrated improvements in both objective knowledge and self-reported confidence following implementation of a subinternship-based bootcamp. However, evidence was limited to a single study and did not assess translation to PGY-1 performance.

PGY-1 Residency Program Based Bootcamp

The final stage for a transition-to-residency surgical bootcamp occurs at the beginning of PGY-1, within the intern's new residency program. These bootcamps aim to ensure incoming interns are well-informed, skilled, and prepared to meet program expectations. Evidence suggests that PGY-1-level bootcamps can rapidly enhance procedural skills and knowledge in surgical interns.[11] The University of Texas Southwestern Medical Center (UTSW) implemented a proficiency-based surgical bootcamp for 20 surgical interns to assess both immediate skill acquisition and long-term retention. Over three days, participants trained in knot-tying techniques (two-handed tie, one-handed tie, surgeons' knot under tension, slip knots) and suturing methods (simple interrupted, horizontal mattress, vertical mattress, interrupted subcuticular, running). Immediate post-tests demonstrated improvements in proficiency based on time and errors. However, performance declined across all tasks at four months post-bootcamp, with greater decreases for more complex techniques ($p = .05$). The authors recommended implementing supplemental skill training throughout the year to improve retention.[12]

The Stanford University Department of Surgery conducted a randomized controlled trial to evaluate their three-day surgical skills bootcamp for incoming PGY-1 interns. Twenty-eight learners were randomized to either the bootcamp intervention (n = 15) or control group (n = 13) and assessed on chest tube insertion, central line placement, peg transfer, and wound closure. Competency evaluations were completed immediately after, one month, and six months post-bootcamp. Similar to the UTSW study, an initial increase in chest tube and central line skill proficiency was seen in the intervention group at baseline and one month later, however this increased performance was largely diminished at six months.[13]

To assess long-term retention, the University of Cincinnati Department of Surgery evaluated whether a structured, longitudinal skills curriculum could enhance surgical skill retention. During orientation, 18 surgical interns completed pre-bootcamp assessments on central line placement, wound closure, knot-tying, and fundamentals of laparoscopic surgery (FLS). Over eight weeks, interns participated in weekly two-hour protected skills sessions, moderated by midlevel residents and evaluated one-on-one by faculty. Post-bootcamp assessments showed improvements in wound closure efficiency, extracorporeal/intracorporeal knot-tying, and peg transfer ($p < 0.01$). Participants had unstructured open access to a skills lab throughout the intern year that housed FLS equipment, suturing kits, and knot tying stations. At the end of the intern year, skills were reassessed, demonstrating maintained proficiency and further improvement in central line and knot-tying skills compared to pre- and post-bootcamp results. Limitations include small sample size and localization to a single institution as well as the interns' knowledge of when the end-of-year test would take place, which the authors mention may have motivated them to practice immediately prior to testing.[14]

The Baystate Medical Center Department of Surgery created a nine-week, simulation-based surgical skills bootcamp for surgical interns, comparing results across four years of intern classes. 30 PGY-1s participated, learning procedural skills such as knot tying, suturing, laparoscopy, chest tube insertion, and central venous catheter placement, as well as engaging in simulated patient care scenarios ranging from shock to surgical emergencies. Pre- and post-bootcamp performance testing showed improvement ($p < 0.001$). Notably, final bootcamp cognitive scores correlated with American Board of Surgery In-Training Examination (ABSITE) scores ($p = 0.01$), suggesting bootcamp performance may predict later clinical knowledge and performance.[15]

Similarly, the University of Florida Department of Surgery evaluated a week-long PGY-1 surgical bootcamp for 15 interns covering knot-tying, suturing, and laparoscopic skills. Participants received personal materials and open access to skill stations throughout the intern year. Performance was assessed in the fall and spring of intern year, as well as fall of PGY-2. Results were compared to a control group of eight PGY-2 residents who received only a brief orientation skills demonstration. By the fall of PGY-2, bootcamp participants outperformed the controls in all surgical skills, including one-hand ties ($p = 0.007$, $p = 0.039$), suturing ($p = 0.001$), clamp-and-ligate efficiency ($p < 0.001$), and laparoscopic skills (peg transfer $p = .018$, circle cut $p = .002$), intracorporeal suturing ($p = .002$). Limitations included single-institution design, small sample sizes, and unreported practice hours throughout PGY-1. However, consistent operative experiences across groups strengthened the credibility of results.[16] Combining a comprehensive surgical bootcamp, self-guided practice opportunities, and long-term assessments resulted in an effective procedure to increase skill acquisition for PGY-1 residents, providing further evidence of the positive benefits of surgical bootcamps.

Across studies, PGY-1 bootcamps were associated with improvements in procedural skills and knowledge, with some evidence demonstrating skill decline over time without adequate reinforcement. Longitudinal training models with protected practice time were associated with improved skill retention.

Hosting bootcamps at the PGY-1 stage allows for specialty-specific skill training, which is often impractical at the MS4 level. An average U.S. medical school class with approximately 150 students typically sees only 7-12 students match into general surgery.[17] When considering students matched into surgical subspecialties, it becomes highly beneficial for a medical school to host a surgical transition-to-residency bootcamp. In contrast, niche specialties, such as dermatology or ophthalmology, often have 2-4 students match per class,[18] making it less feasible for a medical school to invest the time, effort, and finances to implement specialty-specific bootcamps. Thus, such programs are usually seen at the PGY-1 level, where residency programs can tailor specific training to their incoming class.

Comparatively, dermatology receives limited exposure during UME, with most students gaining hands-on experience only through elective or away rotations. The Department of Dermatology at Michigan Medicine implemented a dermatology-specific curriculum for incoming residents to gain objective skillsets and confidence heading into residency. Twenty-four PGY-1 dermatology residents completed an eight-hour preparation course delivered over two days during the first two weeks of residency. Led by faculty volunteers, the curriculum included total body skin exam (TBSE), biopsy set-up, cryotherapy, scraping, electrodesiccation and curettage (EDC), intralesional triamcinolone, and professionalism. Residents completed surveys before, immediately after, and six months following the course. Most participants reported minimal prior experience (< 10 repetitions) with the listed skills. Self-reported

confidence increased across all skills immediately after the course ($p < 0.0001$) and continued to rise at six months ($p < 0.0001$). The majority of residents "agreed" or "strongly agreed" that the bootcamp improved clinical skill competency. Limitations include a small sample size, localization to a single program, and reliance on confidence-based outcomes rather than objective skill performance.[19]

Ophthalmology, another specialty with limited exposure during UME, typically sees meaningful student experience only through electives. Historically, ophthalmology residency begins after a preliminary year in another field, delaying ophthalmology-specific skill acquisition. Although ACGME has mandated integrated intern years for all ophthalmology programs, rapid acquisition of core ophthalmology skills at the onset of residency remains essential. The New York Eye and Ear Infirmary of Mount Sinai implemented a video-based orientation curriculum designed to teach and assess ophthalmic examination skills. Of 11 incoming ophthalmology residents, seven completed a pre-course modified ophthalmic clinical evaluation exercise (OCEX), graded by a senior resident acting as a standardized patient. Using a flipped classroom model, residents viewed short instructional videos on specific exam components, followed by practice sessions the next day. At the end of the two-week orientation, residents were recorded performing the complete ophthalmic exam and scored on a 36-point OCEX. Participants' scores improved from an average of 16.5 to 30.0 ($p < 0.0002$), with anterior segment examinations and applanation showing the greatest gains ($p < 0.01$). Confidence similarly increased from 2.41 to 4.19 ($p < 0.0001$). While this pilot study was limited by small sample size and variability in examiners, results demonstrate the potential of structured bootcamps to accelerate ophthalmic skill development.[20]

At all stages of training, surgical bootcamps were consistently associated with improved confidence and procedural skills. However, most studies were limited by small sample sizes, single-institution design, and focus on short-term outcomes. Evidence for skill retention was strongest for programs incorporating longitudinal training and repeated practice opportunities.

Discussion

This narrative review synthesizes available literature on surgical and specialty-specific bootcamps across multiple stages of medical training, including MS4 home institution, subinternship site, and PGY-1 residency program. The purpose is to provide information on the success and challenges of existing bootcamp programs, as medical institutions around the U.S. move to increase surgical skill training and residency readiness as a result of the ACS joint recommendation.

Bootcamps held at the MS4 home institution benefit both the medical trainee and institution, as a whole. Surgical bootcamps repeatedly demonstrate marked improvement in student clinical knowledge, surgical skill acquisition, and confidence. While not formally measured, participation in the bootcamp may also increase the perceived value of the institution by producing reliable residency-ready graduates. Stony Brook including students from multiple surgical subspecialties of general surgery, orthopedics, integrated plastics, urology, and neurosurgery with utilization of standardized ACS scoring materials allows replication and enhances generalizability.[2] Each of the MS4 studies has limitations, including small sample sizes, single-institution settings, and reliance on subjective confidence measures. Nevertheless, the combination of quantitative and qualitative outcomes, along with consistent findings across institutions, supports MS4 medical school-hosted surgical bootcamps as an effective strategy to ease the transition to surgical residency.

The results from the subinternship hosted bootcamp show hands-on training and early exposure to surgical skills appear beneficial for transitioning to surgical residency. Although evidence is currently limited to a single study, evidence suggests that surgical bootcamps embedded within subinternships can effectively enhance both knowledge and confidence. The lack of broader research highlights a gap in literature and underscores the potential value of subinternship-based bootcamps as interventions that warrant further study across specialties.

PGY-1 surgical bootcamps demonstrate improved resident confidence and procedural skill ability. Evidence from the UTSW, Stanford, and University of Cincinnati bootcamps, highlight that skill acquisition may diminish without reinforcement, underscoring the importance of ongoing, longitudinal training.[12-14] Increasing the frequency of structured skills sessions throughout residency could further enhance retention and mitigate the steep learning curve faced by new surgical trainees. The lack of skill retention results from the UTSW study likely reflect the limited opportunities for PGY-1 residents to apply advanced skills during the first year of surgical residency at

Institution	Learner Level	Specialty	Bootcamp Duration	Sample Size	Skills Covered	Outcome Measures	Main Findings
University of Pennsylvania	MS4	General Surgery	5 days	28 (bootcamp), 9 (control)	Clinical scenarios, chest tube, cricothyroidotomy, NG tube	Confidence (subjective)	Increased confidence; bootcamp group greater than control in cricothyroidotomy and chest tube at 6 months
Stony Brook University & Thomas Jefferson University Hospitals	MS4	General Surgery	2 weeks	12	Patient handoff, suturing, knot tying, central line, chest tube	Skill (objective), confidence (subjective)	Increased ACS skill scores and confidence in all tasks
University of Miami	MS4	General Surgery	4 weeks	28	Consultation scenarios, perioperative care, suturing, knot tying, laparoscopy, ICU procedures	Confidence, preparedness (subjective)	Increased confidence and preparedness for residency
University of Michigan	MS4	General, Thoracic, Vascular Surgery	5 days	23	Cadaver surgical procedures: abdominal, thoracic, vascular	Confidence, anxiety (subjective), procedural knowledge (objective)	Increased confidence and procedural knowledge. Reduced anxiety
Ohio State University	Sub-I	Urology	4 weeks	29	ACGME Level 1 Urology milestones: Foley catheter, bladder irrigation, cystoscopy	Confidence (subjective), knowledge (objective)	Increased confidence, skill mastery, and knowledge
UT Southwestern	PGY-1	General Surgery	3 days	20	Knot tying, suturing	Skill, retention (objective)	Increased proficiency immediate post-bootcamp; decreased performance at four months
Stanford University	PGY-1	General Surgery	3 days	15 (bootcamp), 13 (control)	Chest tube, central line, wound closure, peg transfer	Skill, retention (objective)	Increased proficiency for chest tube and central line at baseline and one month; No significant difference at six months
University of Cincinnati	PGY-1	General Surgery	8 weeks (16 hours)	18	Central line, wound closure, knot tying, laparoscopy	Skill, retention (objective)	Increased proficiency maintained at one year
Baystate Medical Center	PGY-1	General Surgery	9 weeks (36 hours)	30	Patient care scenarios, knot tying, suturing, laparoscopy, chest tube, central venous catheter	Skill, knowledge (objective)	Increased performance; bootcamp scores correlated with ABSITE scores

University of Florida	PGY-1	General Surgery	1 week	15 (bootcamp), 8 (control)	Knot tying, suturing, laparoscopy	Skill, retention (objective)	Bootcamp participants outperformed controls at PGY-2 follow-up
University of Michigan	PGY-1	Dermatology	2 days	24	TBSE, biopsy, cryotherapy, scraping, EDC, injection, professionalism	Confidence (subjective)	Increased confidence immediately after and at six months
NY Ear & Eye Infirmary	PGY-1	Ophthalmology	2 weeks	7	Complete ophthalmic exam	Skill; OCEX (objective), confidence (subjective)	Increased OCEX scores and confidence; greatest gains in anterior segment and applanation skills

Table 1. Bootcamp study characteristics, components, and outcomes.

UTSW, which primarily emphasizes pre- and postoperative care of elective surgical patients.[21] Educational theories such as deliberate practice reinforces the idea that regular, structured practice is imperative to build, maintain, and prevent decay of a skillset.[22]

Several of the reported studies primarily report improvements in self-reported learner confidence and perceived residency preparedness. While increased confidence is important, it may not correspond to improved surgical skill or knowledge. This underscores the importance of combining both subjective and objective assessments when evaluating bootcamp impact.

Future Directions

The responsibility for hosting surgical bootcamps should not fall on a single stage of training. Medical schools, subinternship rotations, and residency programs each provide unique opportunities to strengthen the transition to surgical residency, and a collaborative effort may yield the greatest benefit. Comprehensive training during the final year of medical school can ensure that all students entering surgical specialties acquire core skills, while institution-specific bootcamps at the start of residency can refine techniques and align interns with program expectations. Future research should assess longer-term outcomes such as PGY-1 milestone achievement, procedure logs, resident evaluations, remediation rates, and patient safety metrics.

Challenges and Pitfalls

Despite the benefits demonstrated, widespread implementation of surgical bootcamps faces practical challenges. Establishing simulation-based curricula requires substantial investment in facilities, faculty time, scheduling effort, and institutional resources. Protected schedules must be created to allow learners to participate fully, as demonstrated by the University of Cincinnati's model of pager-free sessions.[14] Residency programs would also need to allocate funding to support intern salaries during bootcamp periods. Furthermore, variability in specialty size may impact feasibility; for example, general surgery typically has larger cohorts of matched MS4s, making a standardized transition course more practical. In contrast, niche fields such as dermatology or ophthalmology, illustrated by smaller match numbers, may be better served by specialty-specific bootcamps at the start of PGY-1.[17,18] These considerations underscore the need for flexible, scalable models of bootcamp design that account for resource availability and specialty-specific needs.

Limitations

This review has its own limitations. Using a narrative design, without following a systematic review protocol, may limit reproducibility and introduce selection bias. The relatively small number of total articles included reflects the specific inclusion criteria and focus on US-based surgical bootcamps targeting MS4 or PGY-1 trainees. Excluding studies outside the US and non-English articles may have omitted relevant international literature. Screening was performed by a

single reviewer, which may increase the risk of selection bias. Publication bias may be present, as studies with positive outcomes are more likely to be published and reporting on these studies may not be indicative of the true effect of bootcamps. The included studies often had differing methodology and small, non-equivalent control groups (e.g., Penn: n = 9 PGY-1, Florida: n = 8 PGY-2), limiting the strength of comparisons and generalizability. These factors should be considered when interpreting the results and recommendations.

Conclusion

Surgical bootcamps can enhance medical students' and residents' self-reported confidence, procedural familiarity, and perceived readiness for training. While most existing studies are limited by small cohorts, single-institution designs, and a focus on short-term outcomes, the overall trend supports their value in bridging the gap between UME and GME. Broader adoption, ideally through coordinated efforts between medical schools and residency programs, could standardize skill acquisition and ease the transition to surgical residency. Future research should prioritize multicenter studies and long-term assessments to determine whether early gains in confidence and surgical skill proficiency translate into sustained performance and improved patient outcomes.

References

1. Joshi P, Hernandez A, Collie BL, et al. Enhancing surgical internship preparedness through a Transition to Internship Bootcamp. Glob Surg Educ. 2024;3(1). doi:10.1007/s44186-024-00234-9
2. Bevilacqua LA, Simon J, Rutigliano D, et al. Surgical boot camp for fourth-year medical students: impact on objective skills and subjective confidence. Surgery. 2020;167(2):298-301. doi:10.1016/j.surg.2019.06.041
3. Ramar K, Oxentenko AS, Dowdy SC. Transforming health care through quality and safety. Mayo Clin Proc. 2025;100(8):1385-1401. doi:10.1016/j.mayocp.2025.04.012
4. Singh P, Aggarwal R, Pucher PH, et al. An immersive "simulation week" enhances clinical performance of incoming surgical interns: improved performance persists at 6 months follow-up. Surgery. 2015;157(3):432-443. doi:10.1016/j.surg.2014.09.024
5. Gervais V, Grabs D, Bougie E, Salib GE, Bortoluzzi P, Tremblay DM. The Montreal Plastic Surgery Residency Bootcamp: structure and utility. Plast Reconstr Surg Glob Open. 2023;11(10):e5337. doi:10.1097/GOX.0000000000005337
6. Galvaing G, Lautrette A, Godet T, et al. Design and evaluation of a surgical boot camp: a step-by-step guide to curriculum development and impact assessment. BMC Med Educ. 2025;25(1):874. doi:10.1186/s12909-025-07545-5
7. Okusanya OT, Kornfield ZN, Reinke CE, et al. The effect and durability of a pregraduation boot camp on the confidence of senior medical students entering surgical residencies. J Surg Educ. 2012;69(4):536-543. doi:10.1016/j.jsurg.2012.04.001
8. Tocco N, Brunsvold M, Kabbani L, et al. Innovation in internship preparation: an operative anatomy course increases senior medical students' knowledge and confidence. Am J Surg. 2013;206(2):269-279. doi:10.1016/j.amjsurg.2012.07.043
9. Thornton SW, Rhodin KE, Leraas HJ, Migaly J, Tracy E, Greenberg JA. The surgical subinternship: a unique opportunity for early leadership development. J Surg Educ. 2023;80(6):757-761. doi:10.1016/j.jsurg.2023.03.006
10. Stout M, Ballinger J, Posid T, et al. Advanced Urology Boot Camp: a simulation-based curriculum to enhance fourth-year medical student procedural competency. Urol Pract. 2023;10(2):196-200. doi:10.1097/UPJ.0000000000000379
11. Blackmore C, Austin J, Lopushinsky SR, et al. Effects of postgraduate medical education "boot camps" on clinical skills, knowledge, and confidence: a meta-analysis. J Grad Med Educ. 2014;6(4):643-652. https://doi.org/10.4300/JGME-D-13-00373.1.
12. Weis JJ, Farr D, Abdelfattah KR, Hogg D, Scott DJ. A proficiency-based surgical boot camp may not provide trainees with a durable foundation in fundamental surgical skills. Am J Surg. 2019;217(2):244-249. doi:10.1016/j.amjsurg.2018.07.040
13. Parent RJ, Plerhoples TA, Long EE, et al. Early, intermediate, and late effects of a surgical skills "boot camp" on an objective structured assessment of technical skills: a randomized controlled study. J Am Coll Surg. 2010;210(6):984-989. doi:10.1016/j.jamcollsurg.2010.03.006
14. Kassam AF, Singer KE, Winer LK, et al. Acquisition and retention of surgical skills taught during intern surgical boot camp. Am J Surg. 2021;221(5):987-992. doi:10.1016/j.amjsurg.2020.09.018
15. Fernandez GL, Page DW, Coe NP, et al. Boot cAMP: educational outcomes after 4 successive years of preparatory simulation-based training at onset of internship. J Surg Educ. 2012;69(2):242-248. doi:10.1016/j.jsurg.2011.08.007
16. Loftus TJ, Filiberto AC, Upchurch GR Jr, et al. Performance improvement with implementation of a surgical skills curriculum. J Surg Educ. 2021;78(2):561-569. doi:10.1016/j.jsurg.2020.08.030
17. Surgery residencies continue to grow in 2025 match. American College of Surgeons. Published March 25, 2025. Accessed September 9, 2025. https://www.facs.org/for-medical-professionals/news-publications/news-and-articles/acs-brief/march-25-2025-issue/surgery-residencies-continue-to-grow-in-2025-match
18. NRMP® celebrates Match Day for the 2024 Main Residency Match®. National Resident Matching Program. Published August 26, 2024. Accessed September 9, 2025. https://www.nrmp.org/about/news/2024/03/nrmp-celebrates-match-day-for-the-2024-main-residency-match-releases-results-for-over-44000-applicants-and-almost-6400-residency-programs
19. Nakamura M, McClean K, Harms K, et al. Dermatology residency preparation curriculum: a model for initiating new residents into dermatology clinical care. Dermatol Online J. 2021;27(12). doi:10.5070/D3271256721
20. Savant SV, Chadha N, Fredrick D, Reddy HS. A novel video-based skills curriculum for incoming ophthalmology residents. Semin Ophthalmol. 2024;39(1):83-88. doi:10.1080/08820538.2023.2239332
21. Training and program structure. UT Southwestern Medical Center. Accessed September 9, 2025. https://www.utsouthwestern.edu/education/medical-school/departments/surgery/education-training/residency-program/program-structure.html
22. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med. 2004;79(10 Suppl):S70-S81. doi:10.1097/00001888-200410001-00022

LETTER TO THE EDITOR

Building a Safer Bridge to Surgical Residency

In reply to Bridging the Gap: A Narrative Review of Surgical and Specialty-Specific Bootcamps in Easing the Transition to Residency

Samer G. Salman, BSA[1*]

[1]School of Medicine, Baylor College of Medicine, Houston, Texas, USA

[*]Corresponding author: Samer G. Salman; Email: samer.salman@bcm.edu

Disclosures: The author reports no financial relationships or conflicts of interest regarding the content herein.

The transition from medical school to surgical residency compresses a profound change in responsibility into a few weeks. New interns begin managing pages, presenting patients, coordinating perioperative care, communicating across teams, and performing procedures under pressure. Their prior clinical exposure and readiness vary widely. National work on the Core Entrustable Professional Activities and surveys of graduating students and surgical program directors have identified persistent gaps in preparation for duties expected at the start of residency.[1, 2] A structured bootcamp can give learners a common starting point and help programs identify areas requiring further coaching before those gaps reach the bedside.

In the accompanying narrative review, "Bridging the Gap: A Narrative Review of Surgical and Specialty-Specific Bootcamps in Easing the Transition to Residency," the authors synthesize 12 US studies spanning fourth-year medical school, surgical subinternship, and the opening of residency. Across specialties and settings, bootcamps were associated with gains in procedural performance, clinical knowledge, confidence, and perceived preparedness, findings consistent with a prior meta-analysis and an objective evaluation of a fourth-year surgical bootcamp.[3, 4] The evidence base remains early, with small cohorts, single-institution designs, and frequent reliance on self-report. The consistency of the findings offers a practical signal: focused preparation during a predictable educational transition can improve readiness.

The review's three-stage framework is especially useful. At the medical school level, a shared foundation can cover handoffs, perioperative assessment, sterile technique, basic suturing, knot tying, line or tube placement, and responses to common clinical scenarios. These activities align closely with competencies expected of learners entering residency.[1] During subinternships, that foundation can be connected to specialty milestones through supervised practice as students assume greater responsibility. Once residency begins, programs can focus on local workflows, escalation pathways, specialty-specific examination skills, and procedural standards. This sequencing reflects the accompanying review's progression from broad preparation to specialty- and institution-specific training. Specialty-focused orientation curricula have produced measurable improvement in clinical examination performance among incoming residents.[5] Coordinating these stages would reduce unnecessary duplication and create deliberate reinforcement.

Durability deserves special attention. The accompanying review contrasts a brief proficiency-based course in which several fundamental surgical skills deteriorated by four months with longitudinal programs that incorporated protected practice, repeated assessment, and continued access to skills laboratories.[6-8] This contrast supports treating bootcamp as the opening phase of a developmental curriculum. A pragmatic cadence to test would pair an intensive orientation course with weekly protected practice during the first eight weeks, followed by brief refresher sessions every two to three months through PGY-1 and formal reassessment near four and twelve months. Spaced practice, direct observation, timely feedback, and progressively complex tasks can carry initial gains into clinical performance. The optimal interval remains uncertain and should be adapted to task complexity, clinical exposure, and learner performance.

Effective implementation can be flexible. The accompanying review identifies practical constraints

including faculty time, protected schedules, simulation facilities, and program size. A systematic review of surgical simulation bootcamps identified delivery at a transition point, curriculum mapping, multimodal instruction, deliberate practice, and formative feedback as central design principles.[9] Scalable programs may combine task trainers, short instructional videos, case discussions, near-peer teaching, faculty observation, and higher-fidelity simulation where available. Learners should receive protected time, clear proficiency criteria, individualized feedback, and an accessible path for additional practice. Readiness also includes recognizing personal limits, escalating concerns early, requesting supervision, and communicating clearly during uncertainty.[1]

Future evaluations should pair confidence measures with objective performance and meaningful workplace outcomes. The highest priorities should be directly observed performance at six and twelve months and clinical transfer, particularly handoffs and responses to deterioration. Multicenter studies could also assess errors, supervisory rescue, and safety events while adjusting for baseline experience and clinical exposure. This addresses the accompanying review's central limitation: immediate improvement is better studied than durability or patient benefit.

Equity should guide access and content. As the review suggests, smaller or less-resourced programs may lack learner volume, faculty time, or simulation capacity. Regional consortia could share videos and assessment checklists, lend portable task trainers or low-cost kits, and rotate quarterly sessions through a central simulation center. This would guarantee access to core materials and minimum practice opportunities while allowing local adaptation. An equitable AI-literacy framework likewise emphasizes accessible, structured curricula.[10]

One unresolved question is who owns continuity as learners move among medical school, subinternship, and residency. We propose testing a learner-controlled competency passport recording practiced tasks, observed proficiency, feedback, and needs for reinforcement. It could help receiving programs target orientation without repeating mastered material or assuming readiness. Its administrative burden and risk as an unintended high-stakes credential require evaluation.

The transition to residency is a shared educational handoff. The accompanying review provides a timely map of where bootcamps can be placed and what they can accomplish. Medical schools, subinternship directors, and residency programs now have an opportunity to build a connected pathway that supports learners and strengthens care for the patients who depend on them. A concrete next step is a one-cycle pilot co-designed by one medical school and a partner residency program, using shared objectives, an orientation bootcamp, a defined refresher schedule, and a six-month assessment. A well-designed bootcamp turns the first day of residency into the next step of guided development.

References

1. Englander R, Flynn T, Call S, et al. Toward defining the foundation of the MD degree: core Entrustable Professional Activities for Entering Residency. Acad Med. 2016;91(10):1352-1358. doi:10.1097/ACM.0000000000001204
2. Lindeman BM, Sacks BC, Lipsett PA. Graduating students' and surgery program directors' views of the Association of American Medical Colleges Core Entrustable Professional Activities for Entering Residency: where are the gaps? J Surg Educ. 2015;72(6). doi:10.1016/j.jsurg.2015.07.005
3. Blackmore C, Austin J, Lopushinsky SR, et al. Effects of postgraduate medical education "boot camps" on clinical skills, knowledge, and confidence: a meta-analysis. J Grad Med Educ. 2014;6(4):643-652. doi:10.4300/JGME-D-13-00373.1
4. Bevilacqua LA, Simon J, Rutigliano D, et al. Surgical boot camp for fourth-year medical students: impact on objective skills and subjective confidence. Surgery. 2020;167(2):298-301. doi:10.1016/j.surg.2019.06.041
5. Savant SV, Chadha N, Fredrick D, Reddy HS. A novel video-based skills curriculum for incoming ophthalmology residents. Semin Ophthalmol. 2024;39(1):83-88. doi:10.1080/08820538.2023.2239332
6. Weis JJ, Farr D, Abdelfattah KR, Hogg D, Scott DJ. A proficiency-based surgical boot camp may not provide trainees with a durable foundation in fundamental surgical skills. Am J Surg. 2019;217(2):244-249. doi:10.1016/j.amjsurg.2018.07.040
7. Kassam AF, Singer KE, Winer LK, et al. Acquisition and retention of surgical skills taught during intern surgical boot camp. Am J Surg. 2021;221(5):987-992. doi:10.1016/j.amjsurg.2020.09.018
8. Loftus TJ, Filiberto AC, Upchurch GR Jr, et al. Performance improvement with implementation of a surgical skills curriculum. J Surg Educ. 2021;78(2):561-569. doi:10.1016/j.jsurg.2020.08.030
9. Young M, Lewis C, Kailavasan M, et al. A systematic review of methodological principles and delivery of surgical simulation bootcamps. Am J Surg. 2022;223(6):1079-1087. doi:10.1016/j.amjsurg.2021.10.044
10. Kumar R, Phadke R, Salman S. Advancing AI literacy in Canadian orthopedic education: a framework for equitable and inclusive training. Can Med Educ J. 2025;16(6):36-38. doi:10.36834/cmej.82734

REVIEW

Ethical Considerations of Gene Therapy in Huntington's Disease: A Review of AMT-130

Kimberly Hoang[1*], Jennifer Escobedo[1], Esther Nwana[2], Faiza Sachwani[3]

[1]Department of Biology and Biochemistry, University of Houston, Houston, Texas, USA
[2]Department of Biomedical Sciences, University of Houston, Houston, Texas, USA
[3]Department of Health and Human Performance, University of Houston, Houston, Texas, USA
[*]Corresponding author: Kimberly Hoang; Email: kkimberlyhoangg@gmail.com
Disclosures: The authors report no financial relationships or conflicts of interest regarding the content herein.

ABSTRACT

Background: Huntington's disease (HD) is a progressive neurodegenerative disorder characterized by motor, cognitive, and psychiatric symptoms, with current treatments focused primarily on symptom management. AMT-130, an experimental gene therapy, targets the underlying genetic cause of HD through adeno-associated virus (AAV)-mediated reduction of Huntingtin protein, offering potential to slow disease progression.

Methods: This study was based on information gathered from free online research databases as well as public search engines. The selected sources address key issues that were identified during preliminary searches into the landscape of HD treatment. The sources used in this paper were peer-reviewed, unless they were part of educational content meant for public awareness and/or personal narratives shared on public websites. The chosen methodology proved to be limiting as access is restricted to uniQure's AMT-130 clinical trial data and other proprietary information. The narrow scope of the review, including the most common topics considered when discussing novel gene therapies: concerns of safety, accessibility, reluctance, and morality, allows for the review to be more focused.

Results: Clinical considerations include the relative safety of AAV vectors, risks associated with high-dose administration, and the procedural challenges of prolonged neurosurgery. Accessibility concerns arise from high treatment costs, limited HD center readiness, and underrepresentation of diverse populations in clinical research. Early intervention with AMT-130 may reduce exposure to medications that cause significant side effects, while psychosocial factors and reluctance to undergo genetic testing influence patient uptake. Social ethical questions also emerge regarding the distinction between essential therapy and elective intervention, particularly in presymptomatic individuals. A major influence to a patient's consent may be the cultural acceptance of human augmentation, particularly through a religious lens. However, a common characteristic across various religions is the approval for the benevolent intent to alleviate human suffering caused by disease, which shifts the ethical question to reevaluating the timing of administration rather than its permissibility.

Conclusion: This review article examines the clinical, ethical, and accessibility implications of AMT-130 in the treatment of Huntington's disease. These considerations underscore the need to balance innovation, risk, and equity in the responsible development and application of gene therapies for HD.

Keywords: Huntington's disease, Ethics, AMT-130, Gene therapy, Accessibility, Clinical risks

Introduction

Huntington's disease (HD) is a progressive neurodegenerative disorder caused by an expanded CAG repeat in the HTT gene and characterized by motor dysfunction, cognitive decline, and psychiatric symptoms.[1-4] Current treatment options primarily focus on symptom management rather than slowing or halting disease progression.[2,3] AMT-130, an experimental gene therapy developed by uniQure, represents a novel approach by targeting the underlying genetic cause of HD through the delivery of an adeno-associated virus (AAV) vector designed to reduce Huntingtin protein expression (Figure 1).[5-8] Its potential to modify disease progression offers hope for patients at a time when no other clinical therapies can slow neurodegeneration.[2,9,10] AMT-130 is currently being evaluated for individuals with early manifest Huntington's disease, typically near or shortly after diagnosis, with the intention of intervening before substantial neurodegeneration occurs, when functional neuronal pathways are still partially preserved.[2,7,10,11] Despite its promise, the rapid development and high cost of AMT-130 have raised critical questions regarding patient accessibility, procedural risks, and the psychosocial impact of treatment, highlighting a gap in ethical considerations that must be addressed as this therapy moves toward clinical use. This review examines emerging clinical evidence alongside ethical and accessibility challenges associated with AMT-130 in Huntington's disease, with emphasis on its therapeutic potential and implementation barriers.

Materials and Methods

The review utilized a narrative literature review approach to examine the clinical and ethical considerations associated with AMT-130 gene therapy for Huntington's disease. Peer-reviewed articles, clinical trial updates, and relevant reports were analyzed to evaluate the therapeutic potential of AMT-130 alongside concerns regarding safety, accessibility, and ethical implementation. The literature was organized into key themes including AAV-based delivery risks, procedural limitation, treatment accessibility, diversity in clinical research, and the psychosocial implications of genetic testing.

To visualize selected sources explaining the AMT-130 therapy process, a flow diagram (Figure 1) was constructed depicting the treatment development following its initiation. To further emphasize the novelty of the gene therapy, current treatment options following HD diagnosis are summarized and compared against AMT-130 in Figure 2. Both figures were assembled in December 2025 using information provided by UniQure at the time of writing, and were created using Lucidchart to supplement reader interpretation of the therapy process and significance.

Results

Safety Concerns and Limitations of AAV

AAVs are generally safe to use as gene therapy vectors.[1,5] A few serotypes are known to possess the desired characteristics for successful gene therapy vectors such as neurotropism and lasting episomal persistence.[1] While AMT 130's clinical trials so far show no adverse effect of the treatment, high-dose administration, like high-dose AMT-130, have raised concerns about potential complications such as hepatotoxicity and dorsal root ganglion inflammation.[1,2] In a clinical trial for a Hemophilia A treatment that used an AAV5 serotype, like AMT-130, 85.8% of participants had elevated levels of the liver enzyme alanine aminotransferase (ALT) after treatment.[12] Of the participants experiencing hepatotoxicity, 12 experienced grade 3 events, meaning they had ALT levels over 520 times the upper limit of the normal range.[12] All participants of the Hemophilia A trial recovered, but that does not take away from the severity of experiencing grade 3 events.[12] If not caught promptly, permanent damage to the liver could occur.[12]

With any viral vector, there's also the possibility for off-target effects such as viral-mediated toxicity, immune responses that restrict the number of treatment administrations, and though the risk of oncogenesis is low, it is not zero and should be taken seriously.[13,14] About 30% of the population may not be eligible for gene therapies that use AAVs because of prior viral exposure/infection.[1,13] Adaptive immune responses developed after such viral contact would deploy antibodies to neutralize the viral capsid proteins which reduce the treatment's efficacy.[1,13]

Procedural Risks of the Treatment

Both the low and high dose of uniQure's AMT-130 treatment require a 12–20-hour complex experimental neurosurgery.[9] AMT-130 targets the caudate nucleus, a structure deep in the brain, that is first affected by HD.6 It is administered via a small catheter into the intracaudate area.[6] Three infusions take place over an 8–10-hour period using MRI to monitor the infusions.[6] AMT-130 then diffuses into the rest of the caudate nucleus, putamen, and eventually the rest of the brain.[6] While there are still many unknowns about patient outcomes of uniQure's trials, there is data to suggest surgical complications double after being under general anesthesia for only 2 hours and further meta-analyses suggest that the risk of complication increases by 14% every 30 minutes past the 2-hour mark.[15,16] Treatment by AMT-130 may be reducing and slowing physiological symptoms, but surgical complications could be a drawback or trade off to this form of gene therapy. Furthermore, prolonged exposure to general anesthesia has been associated with long-term cognitive decline, which directly counteracts the therapeutic goal of preserving cognitive function.[3,16]

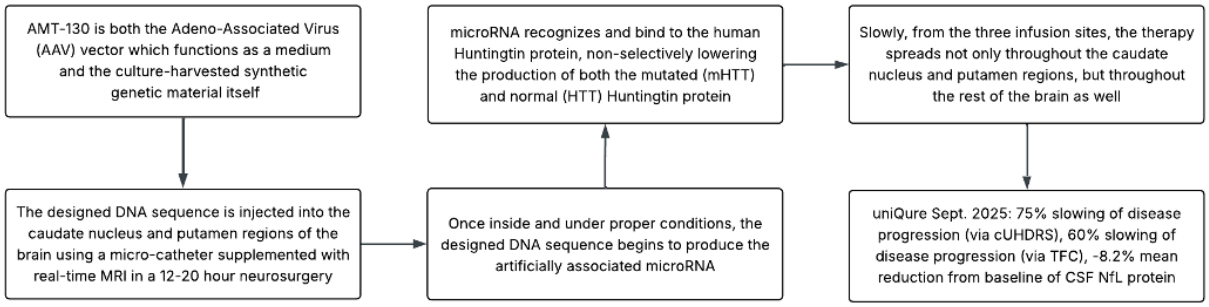

Figure 1. Concept outline of the experimental uniQure AMT-130 gene therapy procedure.

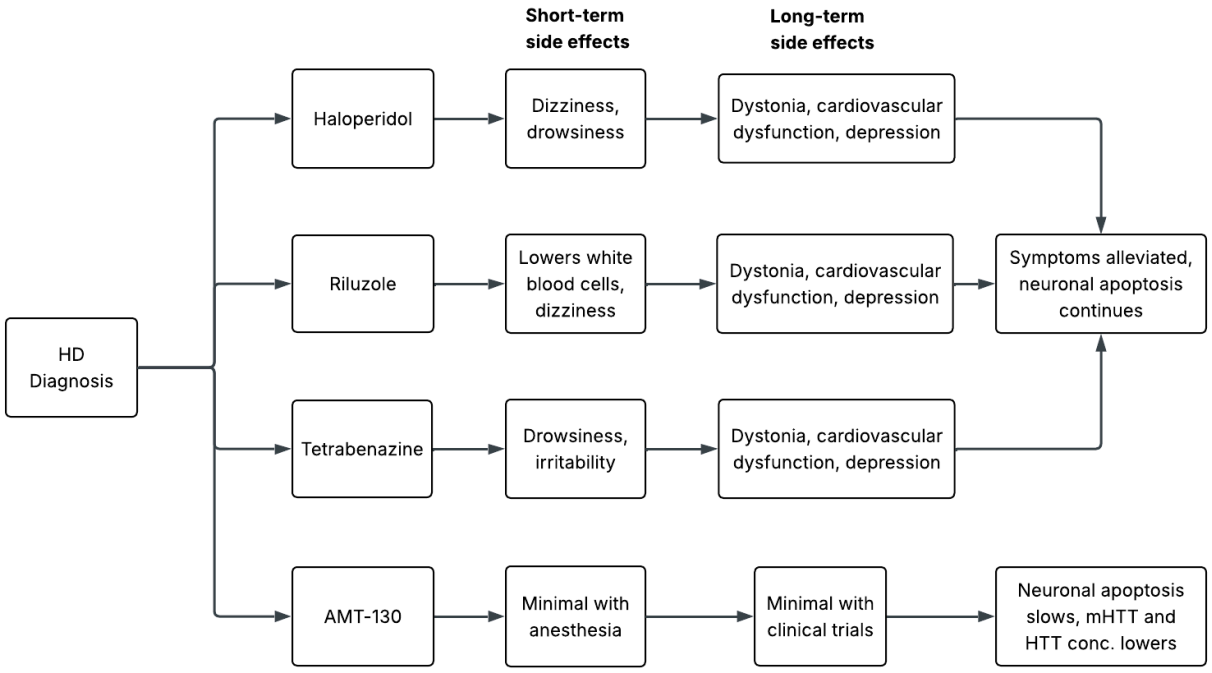

Figure 2. Comparing generalized patient options after HD diagnosis: AMT-130 is demonstrated to be the most effective in alleviating neuronal degradation.

Timeline of Therapy Initiation

At the time of this publication, participants are required to be within the early manifest classification of Huntington's progression to be eligible.[8] However, undiagnosed Huntington's may encourage patients to turn to symptom-mediations such as current drugs used to treat HD-related chorea: haloperidol, riluzole, and tetrabenazine with side effects like dystonia, cardiovascular dysfunction and depression with long-term use.[3,4] This multiple-organ-system impairment interferes with patients' quality of life and may be responded with the increased need for other pharmaceutical interventions.[4] Therefore, initiating the treatment with AMT-130 at the first onset of symptoms following diagnosis is crucial as it has the potential to significantly lessen the length of exposure to comorbidity-inducing chorea medications, making it suitable as a first-line treatment.[4,17] As shown in Figure 2, AMT-130 is currently the only human-tested treatment that modifies the progression of HD, and earlier interventions prove to be more effective in alleviating the symptoms brought on by the disease.[8] Consequently, it is reasonable to consider whether the treatment can be prescribed as a primary therapy.[17] Clinical trials have shown that uniQure's treatment can be effective, but accessibility and procedural risks may prove to be limiting factors in making this a first-line therapy.[8] The high predicted costs and risks associated with the lengthy neurosurgery required to administer AMT-130 are notable concerns.[1,15,16] However, considering the lack of alternatives for sick patients, the risk associated with the treatment is comparatively lower than that of allowing HD to progress.

Accessibility of Treatment

HD gene therapy treatments are costly. Multiple sources predict that AMT-130 will carry a multi-million-dollar pre-insurance cost.[1,7] When low-income individuals cannot always access the healthcare they need, ensuring an equitable roll out and accessibility of AMT-130 is crucial. Tiered pricing has been suggested to minimize access disparities with other AAV therapies, but this system relies on the existence of generic versions of the treatment.[18] Due to FDA scrutiny, approval and widespread rollout of AMT-130 may take many years, and uniQure's patent could delay generic drug production and FDA approval for up to 20 years.[19] Expanded access has been proposed, which would allow individuals with manifest HD access to novel drugs pre-FDA approval.[20]

Additionally, there's the issue of HD centers' readiness to deliver the intracaudate injection that AMT-130 administration requires. Of 40 HD centers that were studied worldwide, 25% of them had the components necessary to perform intracaudate drug injections, but only 6 of the centers could deliver these therapies based on current resources.[21,22] This lack of readiness could mean patients wait approximately 60 months for AMT-130 after 2 years of the therapy becoming available.[21,22] Given the progressive nature of HD, such delays may render many patients ineligible for AMT-130, as symptoms may advance beyond the early manifest stage considered ideal for treatment.

The limited diversity in HD clinical research raises further concerns. With a lack of representation from HD patients of racial and ethnic minorities and pre-symptomatic individuals in clinical trials, concerns arise about AMT-130's efficacy within diverse patients.[22,23] ENROLL-HD's database, which uniQure used as their control group data, has been called into question as the database is 90% White Non-Hispanic.[8,22] While it is true that uniQure's clinical trial demographics seem to align with the demographics of patients diagnosed with HD, it is imperative that there be more minority representation in HD clinical research.[22,24] Black HD patients are often mis- and/or underdiagnosed before being correctly diagnosed despite Black and White Americans having an extremely similar frequency of diagnosis.[25,26] This hinders the understanding of how HD manifests in racial and ethnic minority groups and possibly the efficacy of AMT-130 in diverse patients. Increasing clinical trial participation of minority groups will strengthen the generalizability of studies and is imperative to improve health outcomes. Experts have suggested that fostering better investigator-community relationships, utilizing telehealth tools, and increasing diversity of research staff have positive effect on diversity of clinical trial recruitment.[27]

UniQure reports that most AMT-130 adverse side effects transpire within the first few post-operative weeks, but were addressed to full recovery via either steroids or palliative care.[28] However, participants were strictly monitored over a 5-year duration with quick access to healthcare.[10,28] A longitudinal post-authorization safety and efficacy study would be necessary to ensure long-term safety: additionally accounting for patients who do not have readily access nor the affordability to healthcare that can handle side-effects. In the absence of its wake, mandated risk management plans and systemic checkpoints (e.g., routine pharmacovigilance examinations) have been the

effective rule to mitigating post-market gene therapeutic complications in the European Union.[29]

Testing Reluctance

A large proportion of individuals who have parents with HD choose not to get tested due to the inevitability of the diagnosis.[30] Watching their parents struggle and having no prognosis-modifying options severely discourages this demographic from learning more about their own status.[23,31] Testing for deadly genetic diseases in asymptomatic minors is also discouraged because of the immense psychological distress that the knowledge of their diagnosis could pose.[32] Adults who have tested positive for HD but are presymptomatic express frustration and exhaustion trying to reconcile their current health and experience of life with their fated demise.[33]

AMT-130 offers a disease-modifying option for those who develop HD, conceivably altering the perceived burden that HD places on an individual and their family while alleviating the mental and emotional deterrents for getting tested.[34] Correspondingly, one's access to getting treated with AMT-130 will impact the decision to get tested for HD. The low presymptomatic testing rates have limited the scope of research into the physiological changes that happen prior to the physical and cognitive symptoms that arise during early manifest HD.[23,32]

Cultural Concerns

Medical therapy concentrates on the restoration of health from illness, whereas enhancement aims to supplement beyond a healthy state and is defined as minimal changes from a sustainably normative condition.[35] Ethicists debate the health spectrum parameters, provoking morality arguments on the consequence of human intervention.[36] The growing interest in augmenting human capability challenges the notion that all life is a sacred gift, hence all forms of artificial modification is sacrilegious.[37] Due to AMT-130's introductive nature (unlike the modificative nature of gene editing) the concern of 'playing God' diminishes due to the wide religious acceptance of: alleviating suffering caused by disease, benevolent intent, and preserving the sanctity of life.[37,38] It consequently vastly varies whether gene intervention at germline is considered an essential gene therapy given the life outlook and accounting for anticipation, or considered elective as it is not yet manifest HD.[37] Thus, the culturally ethical question transforms from 'is gene therapy considered immoral?', to 'how early can HD intervention ethically start?'.

Conclusion

AMT-130 represents a significant advancement in the treatment of Huntington's disease, offering the potential to modify disease progression rather than merely manage symptoms. This analysis of uniQure's AMT-130 highlights that its therapeutic promise is constrained by appreciable ethical challenges, including invasive procedural risks, financial barriers and limited clinical infrastructure. Genetic testing avoidance, prolonged uncertainty for presymptomatic individuals, and emotional strain on families are matters that AMT-130 may have a sustained impact on. For the field of gene therapy, these findings emphasize the need to balance scientific innovation with ethical responsibility by ensuring equitable access while minimizing patient harm. Furthermore, our research raises questions that extend beyond the scope of this paper, such as how long-term monitoring of gene therapy patients should be conducted, how early interventions could ethically be initiated in presymptomatic individuals, and how diverse patient populations can be more effectively included in clinical trials. Addressing these questions will be critical for the responsible advancement of gene therapies and the broader application of precision medicine.

References

1. Hao C, Ji H. Gene silencing therapies for Huntington's disease. Neurology Letters. 2025;4:69-75. doi:10.61882/nl.4.1.69
2. Farag M, Tabrizi SJ, Wild EJ. Huntington's disease clinical trials update: October 2025. J Huntingtons Dis. Published online November 26, 2025:18796397251399750. doi:10.1177/18796397251399751
3. Videnovic A. Treatment of Huntington disease. Curr Treat Options Neurol. 2013;15(4):424-438. doi:10.1007/s11940-013-0219-8
4. Huntington's disease. CON-20301692. Published April 25, 2024.
5. Lopez-Gordo E, Chamberlain K, Riyad JM, Kohlbrenner E, Weber T. Natural adeno-associated virus serotypes and engineered adeno-associated virus capsid variants: tropism differences and mechanistic insights. Viruses. 2024;16(3). doi:10.3390/v16030442
6. uniQure. AMT-130 Administration. 2020.
7. Huntington's Disease Society of America. AMT-130: Frequently Asked Questions. Published online November 5, 2025.
8. uniQure. Huntington's disease: silencing the mutant huntingtin gene. Published 2025. Accessed August 5, 2026. https://www.uniqure.com/programs-pipeline/huntingtons-disease
9. A new hope for patients with Huntington's disease? EClinicalMedicine. 2025;88. doi:10.1016/j.eclinm.2025.103612
10. uniQure. uniQure Announces Update on U.S. Phase I/II Clinical Trial of AMT-130 Gene Therapy for the Treatment of Huntington's Disease.
11. uniQure. Huntington's disease: silencing the mutant huntingtin gene. Published online 2025.
12. Zhao Q, Peng H, Ma Y, Yuan H, Jiang H. In vivo applications and toxicities of AAV-based gene therapies in rare diseases. Orphanet J Rare Dis. 2025;20(1). doi:10.1186/s13023-025-03893-z

13. Zaheer A, AboQuella NM, Wadan AHS, et al. CRISPR-based gene therapy for Huntington's disease: current advances and future prospects. Neurogenetics. 2025;26(1):76. doi:10.1007/s10048-025-00858-z
14. Valdmanis PN, Lisowski L, Kay MA. rAAV-mediated tumorigenesis: still unresolved after an AAV assault. Mol Ther. 2012;20(11):2014-2017. doi:10.1038/mt.2012.220
15. Cheng H, Clymer JW, Chen BP, et al. Prolonged operative duration is associated with complications: a systematic review and meta-analysis. J Surg Res. 2018;229:134-144. doi:10.1016/j.jss.2018.03.022
16. Pennings CH, Van Boxtel M, De Korte-De Boer D, Buhre W, Vossen CJ. Anaesthesia as a risk factor for long-term cognitive decline: results of the prospective MAAS cohort study. Eur J Anaesthesiol. 2025;42(5):468-477. doi:10.1097/EJA.0000000000002133
17. Friends of Cancer Research. First-line Therapy. Accessed August 5, 2026. https://friendsofcancerresearch.org/glossary-term/first-line-therapy/
18. Moradpour J, Zhang W, Grootendorst P, Anis AH, Hollis A. Modeling tiered pricing frameworks: a simulation approach. Value Health. 2023;26:351-358. doi:10.1016/j.jval.2022.11.003
19. US Food and Drug Administration. Generic Drugs: Questions & Answers. Published March 16, 2021.
20. US Food and Drug Administration. Expanded Access. Published September 8, 2025. Accessed August 5, 2026. https://www.fda.gov/news-events/public-health-focus/expanded-access
21. Guttman M, Pedrazzoli M, Ponomareva M, et al. The impact of upcoming treatments in Huntington's disease: resource capacity limitations and access to care implications. J Huntingtons Dis. 2021;10(2):303-311. doi:10.3233/JHD-200462
22. Madera A, Schrodt C, Mendizabal A. Disparities in Huntington's disease care and research. Curr Opin Neurol. 2025;38(4). Accessed August 5, 2026. https://journals.lww.com/co-neurology/fulltext/2025/08000/disparities_in_huntington_s_disease_care_and.9.aspx
23. Viau B. To know, or not to know: is the time near to shift genetic testing guidelines? HD Insights. 2025;33:16-19.
24. Mendizabal A, Ogilvie AC, Bordelon Y, Perlman SL, Brown A. Racial disparities in time to Huntington disease diagnosis in North America: an ENROLL-HD analysis. Neurol Clin Pract. 2024;14(5). doi:10.1212/CPJ.0000000000200344
25. Gianattasio KZ, Prather C, Glymour MM, Ciarleglio A, Power MC. Racial disparities and temporal trends in dementia misdiagnosis risk in the United States. Alzheimers Dement (N Y). 2019;5:891-898. doi:10.1016/j.trci.2019.11.008
26. Bruzelius E, Scarpa J, Zhao Y, Basu S, Faghmous JH, Baum A. Huntington's disease in the United States: variation by demographic and socioeconomic factors. Mov Disord. 2019;34(6):858-865. doi:10.1002/mds.27653
27. Kelsey MD, Patrick-Lake B, Abdulai R, et al. Inclusion and diversity in clinical trials: actionable steps to drive lasting change. Contemp Clin Trials. 2022;116. doi:10.1016/j.cct.2022.106740
28. uniQure. uniQure Slides. September 24, 2025.
29. Fritsche E, Elsallab M, Schaden M, Hey SP, Abou-El-Enein M. Post-marketing safety and efficacy surveillance of cell and gene therapies in the EU: a critical review. Cell Gene Ther Insights. 2019;5(11):1505-1521. doi:10.18609/cgti.2019.156
30. Sampaio C, Levey J, Klitzman R. Predictive testing and clinical trials in Huntington's disease: an ethical analysis. Mov Disord. 2018;33(2):243-247. doi:10.1002/mds.27247
31. Caregiver Action Network. Huntington's Disease—What About the Children? Accessed August 5, 2026. https://www.caregiveraction.org/what-about-children/
32. Schmidt HHJ, Barroso F, González-Duarte A, et al. Invited review: management of asymptomatic gene carriers of transthyretin familial amyloid polyneuropathy. Muscle Nerve. 2016;54.
33. Cooper H, Simpson J, Dale M, Eccles FJR. Maintaining psychological well-being when living at risk of Huntington's disease: an interpretative phenomenological analysis. J Genet Couns. 2025;34(2). doi:10.1002/jgc4.1965
34. Guo F, Scholl M, Fuchs EL, Berenson AB, Kuo YF. BRCA testing and testing results among women 18–65 years old. Prev Med Rep. 2022;26. doi:10.1016/j.pmedr.2022.101738
35. Jones DG. Enhancement: are ethicists excessively influenced by baseless speculations? Med Humanit. 2006;32(2):77-81. doi:10.1136/jmh.2005.000234
36. This New Form of Treatment Might Benefit Millions. BMJ Rapid Responses. Accessed August 5, 2026. https://www.bmj.com/cgi/eletters/319/7220/1311#EL2
37. Kalidasan V, Das KT. Playing God? Religious perspectives on manipulating the genome. J Relig Health. 2022;61(4):3192-3218. doi:10.1007/s10943-022-01497-6
38. Bazazzadegan N, Hasanzad M, ASS. Gene therapy and gene editing. In: Hasanzad M, ed. A Glimpse at Medicine in the Future. Springer Nature Singapore; 2024:143-182. doi:10.1007/978-981-97-4974-4_9

REVIEW

Characterizing Hemosiderin in Benign and Malignant Cutaneous Neoplasms: Histopathologic Pearls in Mohs Micrographic Surgery

Alexandra Martinez, BA[1*], Isabella Camacho-Hubbard, MD[1], Nisha Ramani, MD[2], Ida Orengo, MD[1]

[1]Department of Dermatology, Baylor College of Medicine, Houston, Texas, USA
[2]Department of Pathology, Baylor College of Medicine, Houston, Texas, USA
[*]Corresponding author: Alexandra Martinez, BA; Email: alexandra.martinez@bcm.edu
Disclosures: The authors report no financial relationships or conflicts of interest regarding the content herein.

ABSTRACT

Background: Hemosiderin is a frequent histologic finding in Mohs micrographic surgery (MMS) specimens and can appear in both benign and malignant cutaneous conditions. Because MMS surgeons evaluate frozen sections intraoperatively, misinterpretation of hemosiderin can complicate margin assessment and potentially alter surgical decision-making.

Methods: A narrative review of the literature was conducted using PubMed with search terms including "hemosiderin," "Mohs micrographic surgery," "frozen section," "stasis dermatitis," "Kaposi sarcoma," and "atypical fibroxanthoma." Conditions were selected and grouped based on their frequency in Mohs practice and their potential to mimic, obscure, or alter the histologic appearance of cutaneous malignancies.

Results: Hemosiderin deposition was characterized across four conditions commonly encountered by Mohs surgeons: stasis dermatitis, pigmented purpuric dermatoses, Kaposi sarcoma, and pigmented atypical fibroxanthoma. In benign conditions such as stasis dermatitis and pigmented purpuric dermatoses, hemosiderin-laden macrophages can mimic tumor pigment or obscure small tumor nests, risking unnecessary additional Mohs layers. In malignancies such as Kaposi sarcoma, hemosiderin and hemorrhage can overshadow atypical spindle cells, while in pigmented atypical fibroxanthoma, hemosiderin can be mistaken for melanin, creating diagnostic overlap with malignant melanoma. Immunohistochemical stains (e.g., HHV8, S100, Melan A) and special stains (e.g., Prussian blue) were identified as key tools for distinguishing hemosiderin from true tumor pigment.

Conclusion: Recognizing hemosiderin and understanding its histopathologic implications allows Mohs surgeons to avoid both over-treatment and under-treatment. Distinguishing hemosiderin from tumor pigment, correlating clinical context with frozen-section findings, and maintaining a low threshold for ancillary stains or permanent histology when pigment identity is ambiguous can improve intraoperative diagnostic accuracy and surgical precision.

Keywords: Mohs micrographic surgery, Hemosiderin deposition

Introduction

Mohs micrographic surgery (MMS) is a tissue-sparing surgical technique that allows for complete microscopic margin assessment of cutaneous malignancies in real time. While the MMS surgeon acts as the surgeon and pathologist, it is imperative that they recognize histologic artifacts when examining frozen sections under the microscope. One recurring diagnostic challenge is the presence of hemosiderin, an iron-storage pigment formed during the breakdown of red blood cells. Since the human body lacks a mechanism for iron excretion, excess iron accumulates as hemosiderin, which appears on hematoxylin and eosin (H&E) staining as yellow-brown granules. Hemosiderin may be encountered in both benign and malignant processes, and within the context of MMS it can mimic neoplastic pigment, obscure residual tumor, or appear as an artifact related to inflammation or prior biopsy. Misinterpretation may lead to unnecessary additional Mohs layers, incomplete tumor clearance, or confusion regarding tumor type.

Materials and Methods

This narrative review focuses specifically on cutaneous conditions that are both commonly encountered by Mohs surgeons and likely to demonstrate hemosiderin on frozen sections. To identify relevant literature, articles were located on PubMed with combinations of terms such as "hemosiderin," "Mohs micrographic surgery," "frozen section," "stasis dermatitis," "Kaposi sarcoma," and "atypical fibroxanthoma." Studies and reviews were selected for their relevance to histopathologic interpretation during MMS, and the conditions were grouped based on how frequently they appear in Mohs practice and their potential to mimic, obscure, or alter the appearance of cutaneous malignancies. Although hemosiderin has been widely described in dermatopathology, there is currently no cohesive review directed toward MMS surgeons on how this pigment specifically affects intraoperative decision-making. We aim to fill that gap by outlining the clinical and histologic contexts in which hemosiderin appears, explaining how it can alter frozen-section interpretation, and offering practical recommendations to improve diagnostic accuracy.

Results

Hemosiderin in stasis dermatitis (SD), for example, is a benign condition commonly seen when excising skin cancer from the legs in older patients (Figures 1 and 2). SD is caused by chronic venous insufficiency. Increased pressure in the veins leads to cutaneous symptoms such as skin discoloration/hyperpigmentation, itching, burning, redness, and swelling. The hyperpigmentation seen in SD is secondary to hemosiderin deposition.[1] On frozen sections, hemosiderin is often abundant and can mimic brown tumor pigment or obscure small tumor nests.

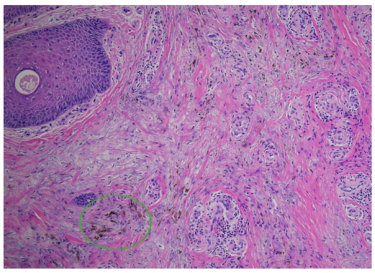

Figure 1. (H&E, 10X): Section shows a lobular pattern of dermal neovascularization with perivascular lymphocytic inflammation, dermal fibrosis and abundant hemosiderin laden macrophages, consistent with stasis dermatitis.

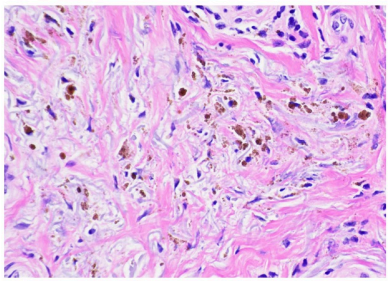

Figure 2. (H&E, 40X): High magnification of the image showing a golden brown refractile pigment presenting as irregular "chunky" aggregates, consistent with hemosiderin pigment. Presence of hemosiderin pigment demonstrates previous hemorrhage.

In pigmented basal cell carcinoma, hemosiderin may resemble melanin, complicating the distinction between tumor-related pigment and background inflammatory pigment. Surgeons may confuse hemosiderin-laden macrophages with residual tumor cells, which can lead to unnecessary Mohs stages.[1,2]

Pigmented purpuric dermatoses (PPD) is another benign, cutaneous condition that can present with hemosiderin deposition. PPDs are chronic, benign purpuric skin lesions that occur most commonly on the lower extremities, similarly to SD. The exact etiology of PPDs is unknown, but they are believed to be caused by capillary dysfunction that leads to hemosiderin deposition in the superficial dermis.[2] During MMS, these deposits can be misinterpreted as tumor-associated pigment, particularly in cases where melanoma or pigmented basal cell carcinoma is suspected. The accompanying inflammation can also obscure small islands of tumor, making careful evaluation critical.[2]

In addition to presenting in benign conditions, hemosiderin can also occur histologically in cutaneous malignancies, like Kaposi Sarcoma (KS). KS can present clinically similar to SD, especially on the lower legs. Microscopic examination of KS show poorly

formed slit-like vascular channels dissecting dermal collagen with bland appearing spindle cells, extravasated red blood cells, and hemosiderin deposition (Figure 3). These features can be subtle on frozen sections, especially in early lesions, and the presence of hemosiderin and hemorrhage may overshadow atypical spindle cells. Typically, immunostaining for HHV8 shows diffuse nuclear reactivity, confirming the morphologic impression of KS (Figure 4).[3]

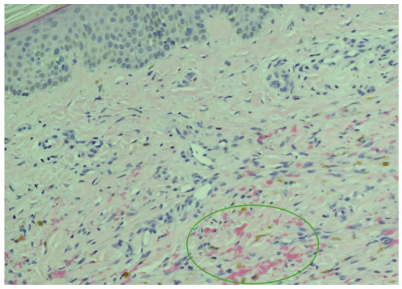

Figure 3. (H&E, 10X): Section from the lower leg of an elderly patient shows bland appearing spindle cells, extravasated red blood cells and prominent hemosiderin deposition.

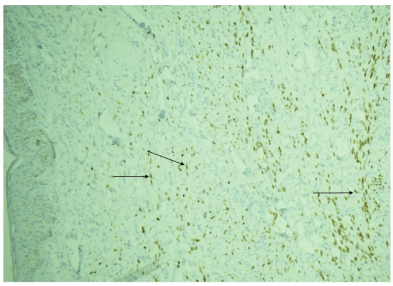

Figure 4. (HHV8 Immunostain, 10X): The immunostain for HHV8 shows the spindle cells with diffuse nuclear reactivity, confirming the diagnosis of Kaposi Sarcoma.

Pigmented atypical fibroxanthoma (AFX) is another cutaneous neoplasm with hemosiderin deposition that can be mistaken both clinically and histologically for malignant melanoma.[4] It is a rare, mesenchymal tumor that typically appears on sun damaged areas of older patients and presents as an erythematous or hyperpigmented plaque that may bleed, ulcerate, or crust over. It can present with a range of histologic features such as atypical histiocytes, fibroblasts, monomorphous spindle cells, clear cells, or granular cells thus making diagnosis difficult.[4] Melanoma tends to have a slower evolution, more variable pigmentation, and the classic ABCDE features (asymmetry, border irregularity, color variation, diameter >6 mm, and evolution). AFX lacks the jet-black or blue-black pigmentation typical of many melanomas and usually has a more friable, exophytic surface rather than the smooth or irregular pigmented plaque/nodule seen in melanoma. Despite these distinctions, overlap can occur, and histologic differentiation is essential.[4,5]

Histologically, AFX can mimic malignant melanoma, especially when hemosiderin resembles melanin on H&E staining. In AFXs the hemosiderin granules in the cytoplasm can be identified with Prussian blue staining.[4, 5] In contrast, the pigmentation in melanomas is due to melanin production by the tumor cells that can be identified with immunohistochemical stains specific for melanoma such as S100 and Melan A.[4, 5] Misinterpretation during MMS may lead to inappropriate margins or inaccurate tumor classification, underscoring the importance of correlating histologic, immunohistochemical, and clinical features.

Discussion

Across these conditions, misinterpretation of hemosiderin can significantly impact Mohs surgeons' intraoperative decision-making. Hemosiderin-laden macrophages may be mistaken for residual tumor and prompt unnecessary additional layers. However, dense pigment or hemorrhage can obscure subtle tumor islands or atypical spindle cells and increase the risk of incomplete excision. Hemosiderin may also mimic melanin, leading to diagnostic confusion between AFX and melanoma or between pigmented inflammatory lesions and true pigmented neoplasms. In vascular proliferations, hemorrhage and hemosiderin can mask abnormal endothelial or spindle cell growth, delaying recognition of malignancy.

To minimize errors, Mohs surgeons benefit from integrating clinical context with frozen- section interpretation. Pigmentation limited to chronically dependent lower legs suggests benign processes such as stasis dermatitis or PPD. When pigment identity is uncertain, rapid stains such as Prussian blue can help distinguish hemosiderin from melanin. If features suggest a vascular neoplasm such as KS, deeper frozen cuts or referral to permanent sections may be warranted, especially because immunostaining is necessary for definitive diagnosis. Documenting ambiguous areas, taking additional sections, and maintaining a low threshold for permanent histology can improve diagnostic accuracy.

Conclusion

Recognizing hemosiderin and understanding its implications allows Mohs surgeons to avoid both over-treatment and under-treatment. By distinguishing hemosiderin from tumor pigment, interpreting inflammatory or vascular changes appropriately, and correlating clinical and microscopic findings, surgeons can optimize margin assessment and surgical precision. Although hemosiderin is not a defining feature of cutaneous malignancy, its presence provides important clues about prior hemorrhage, chronic inflammation, or underlying pathology. A thorough understanding of its histopathologic significance enhances intraoperative accuracy and contributes to safer, more effective skin cancer treatment.

References

1. Northcutt AD, Tschen JA. The routine use of iron stain for biopsies of dermatoses of the legs. J Cutan Pathol. 2015;42(10):717-721. 10.1111/cup.12319.
2. Rzepecki AK, Blasiak RC. Stasis dermatitis: differentiation from other common causes of lower leg inflammation and management strategies. Curr Geriatr Rep. 2018;7:222-227. doi:10.1007/s13670-018-0257-x.
3. Patel RM, Goldblum JR, Hsi ED. Immunohistochemical detection of human herpesvirus 8 latent nuclear antigen-1 is useful in the diagnosis of Kaposi sarcoma. Mod Pathol. 2004;17(4):456-460. doi:10.1038/modpathol.3800075.
4. Diaz-Cascajo C, Weyers W, Borghi S. Pigmented atypical fibroxanthoma: a tumor that may be easily mistaken for malignant melanoma. Am J Dermatopathol. 2003;25(1):1-5. doi:10.1097/00000372-200302000-00001.
5. Zhu TH, Painter C, Hendrix JD. Hemosiderotic (pigmented) atypical fibroxanthoma mimicking malignant melanoma. Dermatol Online J. 2022;28(6). doi:10.5070/D328659728.

RESEARCH LETTER

State-level Correlation of Binge Drinking and Stimulant and Alcohol Co-use Mortality in the United States, 2012-2022

Geoffrey Zhang, BS[1*]

[1]School of Medicine, Baylor College of Medicine, Houston, Texas, USA
[*]Corresponding author: Geoffrey Zhang, BS; Email: geoffrey.zhang@bcm.edu
Disclosures: The author reports no financial relationships or conflicts of interest regarding the content herein.

ABSTRACT

Background: Excessive alcohol consumption and stimulant-related overdoses both contribute substantially to U.S. mortality. However, little is known about whether state-level binge drinking prevalence is associated with stimulant–alcohol co-use mortality.

Methods: We conducted a cross-sectional analysis using state-level data (N = 51; 50 states and the District of Columbia) from the Behavioral Risk Factor Surveillance System (BRFSS) and CDC WONDER for 2012–2022. Binge drinking prevalence was averaged across years. Associations were assessed with Pearson correlation coefficients, with sensitivity analyses excluding the District of Columbia.

Results: State-level binge drinking prevalence was significantly correlated with alcohol–cocaine mortality ($r = 0.247$, $p < .001$). This relationship persisted, though weaker, after excluding the District of Columbia ($r = 0.140$, $p = .024$). No significant correlation was observed between binge drinking prevalence and alcohol–psychostimulant mortality ($r = –0.055$, $p = .518$).

Conclusion: Higher state-level binge drinking prevalence was associated with greater mortality involving alcohol and cocaine co-use, but not alcohol and psychostimulants co-use. These findings suggest the need for integrated alcohol prevention and stimulant harm reduction strategies, particularly in states with high binge drinking rates. Further research is warranted to explore causal pathways using individual-level data.

Keywords: Binge drinking, Cocaine, Psychostimulants, Alcohol, Co-use, Mortality

Introduction

Excessive alcohol consumption remains a major and preventable cause of mortality in the United States, contributing to a substantial public health burden.[1] Binge drinking is a particularly high- risk behavior and a significant driver of alcohol-related harms.[2] Concurrently, the United States is grappling with a continuing stimulant overdose crisis.[3,4]

The co-use of alcohol and stimulants presents a complex and dangerous public health challenge, as the interaction between these substances can significantly amplify health risks. For instance, the concurrent use of alcohol and cocaine leads to the metabolic formation of cocaethylene, a substance with a longer plasma half-life and greater cardiotoxic effects than cocaine alone.[5-8] While the pharmacological risks are well-documented, less is understood about the relationship between population-level alcohol use patterns and mortality involving the co-use of alcohol and stimulants.

Understanding whether state-level variations in high-risk drinking behaviors are associated with mortality rates from alcohol and stimulant use could help identify areas where integrated prevention and harm reduction strategies are most needed. This study aims to investigate the correlation between the state-level prevalence of binge drinking and mortality rates involving the co-use of alcohol with cocaine and other psychostimulants across the United States.

Methods

This cross-sectional study utilized publicly available, aggregated data and was thus exempt from IRB review.

Data and Measures

Average annual state-level prevalence of binge drinking for adults aged 18 and older was calculated using data from the 2012–2022 Behavioral Risk Factor Surveillance System (BRFSS).[9] Binge drinking was defined as self-reporting five or more drinks on one occasion for men or four or more for women within the last 30 days.

Age-adjusted mortality rates for deaths involving the co-use of stimulants and alcohol were obtained from the CDC WONDER Multiple Cause of Death database for the years 2012–2022.[10] Overdose deaths were identified by the International Classification of Diseases, Tenth Revision (ICD-10) underlying cause-of-death codes X40–X44, X60–X64, X85, and Y10–Y14. We analyzed two categories of co-use deaths using contributing causes: (1) deaths with codes for cocaine (T40.5) and alcohol (T51), and (2) deaths with codes for other psychostimulants with abuse potential (T43.6) and alcohol (T51).

Statistical Analysis

The association between state-level binge drinking prevalence and the two mortality rates was assessed using Pearson correlation coefficients. A sensitivity analysis was performed after excluding the District of Columbia. A p-value of less than .05 was considered significant. These analyses describe associations at the population level and do not support a direct individual-level relationship between binge drinking and co-use mortality. Analyses were performed using R, version 4.4.2.[11]

Results

The primary analysis demonstrated that states with a higher prevalence of binge drinking also tended to have higher mortality rates from the co-use of cocaine and alcohol. This positive correlation was statistically significant (r = 0.247, p < .001) (Figure 1). This association persisted, though weakened, after the District of Columbia was excluded in a sensitivity analysis (r = 0.140, p = .024).

For other psychostimulants, however, no such relationship was observed. The correlation between state-level binge drinking prevalence and mortality involving other psychostimulants and alcohol was not statistically significant (r = -0.055, p = .518).

Discussion

Our results identify an ecological link between binge drinking and cocaine-alcohol mortality. There was a significant, though modest, positive correlation between state-level binge drinking prevalence and mortality involving cocaine and alcohol co-use (r = 0.247). In contrast, no such relationship was observed for deaths involving other psychostimulants and alcohol.

The association with cocaine may be explained by both pharmacological and behavioral factors. The co-ingestion of alcohol and cocaine uniquely produces cocaethylene, a metabolite with heightened cardiotoxicity and a longer half-life than cocaine alone.[12,5] Furthermore, social settings that facilitate binge drinking may also promote concurrent cocaine use, increasing population-level risk. The absence of a similar correlation for other psychostimulants, such as methamphetamine, could suggest different contexts of use or other primary drivers including differing geographic and demographic patterns of use.

Conclusion

These findings suggest that public health strategies should consider integrating alcohol misuse prevention with stimulant harm reduction, particularly in states with high binge drinking rates and cocaine-involved mortality. Clinicians should be aware of this amplified

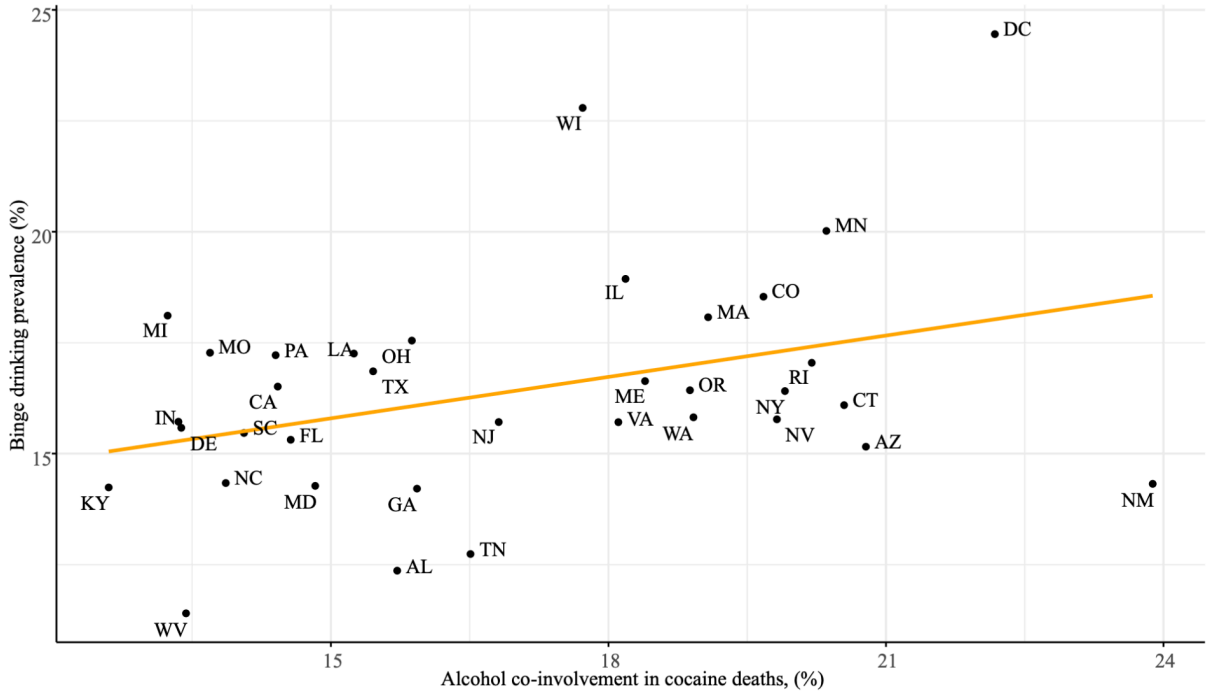

Figure 1. State-Level Binge Drinking Prevalence vs. Alcohol-Cocaine Co-Use Mortality Rate (per 100,000), 2012-2022.

risk and screen for co-use.

This study has several limitations. These findings describe population-level associations and do not support a direct individual-level relationship. Self-reported binge drinking data can be affected by recall bias. Underreporting of substances on death certificates may affect results. Analysis did not control for state-level confounders such as poverty, urbanicity, and drug supply characteristics, including the presence of fentanyl-contaminated cocaine.

Future research using individual-level data is needed to explore the causal pathways behind this association.

References

1. White AM, Castle IP, Hingson RW, Powell PA. Using death certificates to explore changes in alcohol-related mortality in the United States, 1999 to 2017. Alcohol Clin Exp Res. 2020;44(1):178-187. doi:10.1111/acer.14239

2. National Institute on Alcohol Abuse and Alcoholism. Understanding binge drinking. National Institutes of Health. Accessed July 15, 2026. https://www.niaaa.nih.gov/publications/brochures-and-fact-sheets/binge-drinking

3. Kariisa M, Scholl L, Wilson N, Seth P, Hoots B. Drug overdose deaths involving cocaine and psychostimulants with abuse potential—United States, 2003-2017. MMWR Morb Mortal Wkly Rep. 2019;68(17):388-395. doi:10.15585/mmwr.mm6817a3

4. Spencer MR, Miniño AM, Warner M. Drug Overdose Deaths in the United States, 2001-2021. NCHS Data Brief No. 457. National Center for Health Statistics; 2022. doi:10.15620/cdc:122556

5. Pennings EJ, Leccese AP, Wolff FA. Effects of concurrent use of alcohol and cocaine. Addiction. 2002;97(7):773-783. doi:10.1046/j.1360-0443.2002.00158.x

6. Havakuk O, Rezkalla SH, Kloner RA. The cardiovascular effects of cocaine. J Am Coll Cardiol. 2017;70(1):101-113. doi:10.1016/j.jacc.2017.05.014

7. van Amsterdam J, Gresnigt F, van den Brink W. Cardiovascular risks of simultaneous use of alcohol and cocaine: a systematic review. J Clin Med. 2024;13(5):1475. doi:10.3390/jcm13051475

8. Rash CJ, Alessi SM, Petry NM. Cocaine abusers with and without alcohol dependence respond equally well to contingency management treatments. Exp Clin Psychopharmacol. 2008;16(4):275-281. doi:10.1037/a0012787

9. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System Survey Data, 2012-2022. U.S. Department of Health and Human Services; 2023.

10. National Center for Health Statistics. About Multiple Cause of Death, 2011-2022. CDC WONDER Online Database. Centers for Disease Control and Prevention. Released 2025. Accessed May 15, 2025.

11. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2024. https://www.R-project.org/
12. Andrews P. Cocaethylene toxicity. J Addict Dis. 1997;16(3):75-84. doi:10.1300/J069v16n03_08

Supplemental Issue
Fall 2026

ABSTRACT

Undergraduate Medical Humanities as a Pedagogical Framework for Early Professional Identity Formation

Rachael M. Johnson, BA[1]*, Michael J. Pottash, MD, MPH[2], Julia F. Langley, MA, MBA[3]

[1]School of Medicine, Baylor College of Medicine, Houston, Texas
[2]MedStar Washington Hospital Center, Georgetown University School of Medicine, Washington, D.C., USA
[3]Georgetown Lombardi Comprehensive Cancer Center, Georgetown University School of Medicine, Washington, D.C., USA
*Corresponding author: Rachael M. Johnson, BA; Email: rachael.johnson@bcm.edu
Disclosures: The authors report no financial relationships or conflicts of interest regarding the content herein.

Keywords: Medical humanities, Medical education, Undergraduate education, Premedical curriculum, Professional identity, Reflection

Introduction
Physician identity formation is a complex, ongoing process that begins long before students enter the clinical training phase of medical education. While the traditional pre-medical curriculum often emphasizes the biological sciences, it may overlook the humanistic and social dimensions of healthcare. In response to this gap, we developed a semester-long undergraduate course, *Introduction to the Medical Humanities*, at Georgetown University. The course aimed to supplement the pre-medical curriculum by introducing students to patient and public healthcare narratives, thereby re-centering the human experience in their preparation for a medical career. Findings reflect the experiences of 25 students from a single course at one university and therefore may not be broadly generalizable. Nonetheless, this study provides valuable insight into how premedical humanities education may impact students' understandings of health, illness, and their developing sense of professional identity.

Methods
The course enrolled 25 undergraduate students. Educational methods included a combination of pre-class readings, reflective writing assignments, and small-group discussions. Readings were drawn from literature, memoir, journalism, and scholarly texts, with topics spanning chronic illness, disability, caregiving, cultural representations of medicine, and structural determinants of health. Each class began with students sharing their reflections, which served as a springboard for dialogue. At the conclusion of the semester, students completed a final reflective writing assignment in which they were asked to synthesize their learning and describe how the course influenced their perception of medicine.

Results
Analysis of students' final reflections revealed six major themes: (1) students gained insight into how people's unique experiences, values, and perspectives influence the way they interact with the healthcare system, and (2) cultivated empathy through exposure to new ideas that challenged previously held beliefs regarding disabling conditions such as injury and illness. Furthermore, (3) the practice of vulnerability helped students build community and learn from their peers. Students also (4) learned how the humanities, and its interdisciplinary approach, inform the study of health and medicine, (5) developed an understanding of how language and culture can shape interpersonal dynamics in the clinical setting, and (6) realized that discomfort is often an essential aspect of emotional and intellectual growth.

Conclusion
This study suggests that undergraduate medical humanities education can meaningfully contribute to the early stages of physician identity formation. Through reflection, discussion, and engagement with diverse narratives, students not only gained a more

nuanced understanding of the human experience of illness but also began to develop key interpersonal and ethical competencies. Students demonstrated increased awareness of how culture, language, disability, and lived experience can shape interactions with the healthcare system, while also recognizing the importance of vulnerability, active listening, and self-reflection in patient care. Furthermore, many students described becoming more comfortable engaging with ambiguity and emotional discomfort, both of which are central to the practice of medicine. Medical educators may benefit from integrating similar approaches to develop empathy, cultural sensitivity, and critical self-awareness among aspiring physicians.