

Recent Advances in Implemented Medical Education Across Different Medical Specialties: A Meta-Analytic Study

Abdalmohsen M. Alomair

Medical Education Department, College of Medicine, King Faisal University, Al Ahsa, Saudi Arabia

Received: 12 July, 2026; Revised: 20th July 2026; Accepted: 8th Aug, 2026; Available Online: 31, Aug, 2026

ABSTRACT

Background: Medical education witnessed rapid transformation over the past decade, driven by simulation technology, extended reality (virtual and augmented reality), competency-based frameworks, interprofessional education (IPE), digital and blended learning, and, most recently, generative artificial intelligence (AI). Whether these innovations translate consistently into measurable clinical and educational benefit across specialties remains incompletely quantified.

Objective: To synthesize quantitative evidence from systematic reviews and meta-analyses published between 2010 and 2025 on the effectiveness of contemporary medical education interventions, and to evaluate their translational application across surgical, medical, nursing, and allied clinical specialties.

Methods: A systematic search of PubMed/MEDLINE, Embase, Scopus, Web of Science, Cochrane CENTRAL, and ERIC was conducted for meta-analyses and systematic reviews reporting quantitative effect estimates for medical education interventions. Sixty-one meta-analytic syntheses meeting eligibility criteria were retained following PRISMA 2020 methodology. Standardized mean differences (SMD) with 95% confidence intervals were extracted and pooled narratively across six domains: simulation-based education, extended reality (VR/AR), competency-based medical education (CBME), interprofessional education, e-learning/blended learning, and AI-augmented education.

Results: Simulation-based education demonstrated a large, consistent effect on clinical skill acquisition (pooled SMD range 0.70–0.86). Extended reality training produced the largest effect sizes among all domains, particularly for surgical and procedural competencies (SMD 1.08–1.75). Competency-based education interventions showed moderate-to-large effects on research and clinical skills (SMD $\approx$ 0.69). Interprofessional education demonstrated smaller but clinically meaningful effects on patient-reported outcomes (SMD 0.18–0.32). Blended and e-learning approaches performed two comparable tools. Several analyses exceeded traditional face-to-face instruction. Emerging evidence on AI-augmented and learning-analytics-driven education is promising but heterogeneous and still maturing.

Conclusion: Technology-enhanced and competency-oriented educational innovations produce moderate-to-large, reproducible improvements in knowledge, skill, and to a lesser extent, patient-level outcomes, with the strongest and most consistent translational uptake in surgical and procedural specialties. Interprofessional and AI-driven approaches show favorable but less mature evidence, warranting continued rigorous evaluation.

Keywords: medical education; meta-analysis; simulation-based training; virtual reality; competency-based medical education; interprofessional education; e-learning; artificial intelligence; translational education; Vancouver referencing

How to cite this article: Abdalmohsen M. Alomair, Recent Advances in Implemented Medical Education Across Different Medical Specialties: A Meta-Analytic Study. *Int J Drug Deliv Technol.* 2026;16(83s): 474-485; DOI: 10.25258/ijddt.16.83s.60

Source of support: Nil.

Conflict of interest: None

INTRODUCTION

Medical education is currently witnessing a period of rapidly accelerated methodological change.⁽¹⁾ The transition from time-based, apprenticeship-oriented curricula toward competency-based medical education (CBME) reoriented the field toward measurable outcomes rather than fixed durations of training.⁽²⁾ Improved innovation, in medical education entailed the usage high- and low-fidelity simulation, virtual and augmented reality (VR/AR), e-learning platforms as well as the usage of generative artificial intelligence (AI). All these modern tools have tremendously added and reshaped many clinical knowledge and procedural competency acquisition and assessment within the field of medical education.^(3–5)

Despite an expanding evidence base, controversy exists as regards; the comparative magnitude of effect across different pedagogical innovations. This includes the fact that whether gains observed in controlled educational settings would be translated into durable clinical performance and/or this translational uptake is uniform across medical specialties or concentrated within procedurally intensive fields such as surgery and anesthesiology.^(6,7) Continuing medical education (CME) and continuing professional development (CPD) strategies, for instance, have been studied for over seventy years, yet syntheses of their comparative effectiveness remain fragmented across specialty-specific literatures.⁽³⁾

This meta-analytic study aimed to address these gaps by systematically aggregating quantitative findings by retrieving data from meta-analyses and systematic reviews published between 2010 and 2025, spanning six major domains of contemporary medical education innovation; simulation-based education, extended reality (VR/AR) training, competency-based medical education, interprofessional education (IPE), e-learning and blended learning as well as AI- and learning-analytics-augmented education. The primary aim is to quantify pooled effect sizes within each domain and to map the extent and pattern of translational application across different disciplines including; surgery, internal medicine, emergency medicine, nursing, anesthesiology/critical care, primary care, and pediatric specialties.

Currently, the existing meta-analyses tend to be confined to a single modality (e.g., simulation or VR) or a single specialty (e.g., orthopedic surgery, or diabetes care in primary practice).^(8–13) This study offers educators, curriculum designers, and health-system administrators an evidence-based hierarchy of intervention effectiveness to inform resource allocation. This is performed by consolidating effect estimates across domains and specialties.

This study was conducted with three specific objectives; to quantify and compare pooled standardized effect sizes across the six identified domains of medical education innovation, to characterize the pattern of translational application of each domain across medical and surgical specialties as well as identification of the methodological gaps and evidence limitations that should guide future primary research and higher-order synthesis.

METHODS

This meta-analytic study was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement.

Electronic databases searched included PubMed/MEDLINE, Embase, Scopus, Web of Science, the Cochrane Central Register of Controlled Trials (CENTRAL), and ERIC, from database inception through July 2025. Search terms combined controlled vocabulary (MeSH) and free-text keywords across four concept blocks; medical/health professions education, named pedagogical innovations (simulation, virtual reality "VR", augmented reality "AR", competency-based education, interprofessional education, e-learning, blended learning, artificial intelligence, learning analytics), study design filters for systematic review and meta-analysis. Outcome terms (knowledge, skill, competence, patient outcome, translational) were also included in the search. The full electronic search strategy is reproduced in Annexure B.

Data were retrieved from eligible studies if they were; systematic reviews or meta-analyses, published in English between 2010 and 2025 (with selected foundational studies from earlier years retained where

cited seminally). A reported quantitative pooled effect estimates (standardized mean difference, odds ratio, or equivalent) for an educational intervention within medicine, surgery, nursing, or allied clinical training, and specified the clinical specialty or training population.⁽³⁾ Reviews without extractable quantitative synthesis, non-health-professions education studies, and duplicate/overlapping cohorts were excluded.

Two-stage screening (title/abstract, then full text) was performed against the above-mentioned eligibility criteria. Data extracted from each study included meta-analysis comprised: first author and year, domain of intervention, comparator, outcome type (knowledge, skill/psychomotor, attitude, patient outcome), pooled effect size with 95% confidence interval, number of primary studies and participants pooled, and reported clinical specialty. The full extraction matrix is presented in Annexure C.

Methodological quality of included meta-analyses was appraised using "AMSTAR-2" A MeaSurement Tool to Assess systematic Reviews domains, and heterogeneity was qualitatively classified as low, moderate, or high based on reported I^2 statistics where available. A summary of the quality matrix is provided in Annexure D.

Because included studies varied substantially in outcome metrics, populations, and comparators, a formal re-pooled random-effects meta-analysis of primary data was not undertaken. However, this study performed a structured meta-analytic synthesis of already-pooled effect estimates (a "meta-analysis of meta-analyses" / umbrella-style synthesis). This method is increasingly used when primary-study heterogeneity precludes direct re-analysis. Effect sizes were reported as standardized mean differences "SMD" with 95% confidence intervals. Yet, the original review reported Cohen's d or Hedges' g , values were retained as reported and treated as equivalent to SMD for cross-domain comparison purposes. Weighted domain-level means (Figure 2) were calculated by weighting each contributing meta-analysis via its reported number of pooled primary studies.

As the included studies varied substantially in outcome metrics, populations, and comparators, a formal re-pooled random-effects meta-analysis of primary data was not undertaken. Effect sizes were reported as standardized mean differences "SMD" with 95% confidence intervals.

RESULTS

A total of 4,812 records were identified through database searching. After de-duplication, screening, and full-text review, 61 meta-analyses/systematic reviews with quantitative synthesis met the current study inclusion criteria (Annexure A, Figure A1). Involved reviews spanned publication years 2011–2025 and collectively synthesized findings from several thousand primary studies and hundreds of thousands of learners across specialties.

Table 1 showed summarizes for the characteristics of the primary meta-analyses underpinning each domain of this synthesis. The full extraction matrix, including

secondary and supporting reviews, is presented in Annexure C.

Table 1. Characteristics of Key Meta-Analyses Included in the Synthesis, by Educational Domain

Domain	Key Source(s)	Primary Studies (n)	Pooled Effect (SMD/95% CI)	Principal Outcome
Simulation-based education	McGaghie et al., 2011 (8)	14	0.71 (0.65–0.76)	Procedural skill acquisition
Simulation-based education (fidelity-stratified)	Kim et al., 2016 (9)	40	0.70 (pooled random effects)	Cognitive, psychomotor, affective domains
Technology-enhanced simulation	Cook et al., 2011 (14)	609	Large, heterogeneous	Knowledge, skill, behaviors, patient effects
Extended reality (VR/AR) – orthopedics	Li et al., 2025 (10)	23	1.08–1.75	Knowledge, clinical operation, surgical design
Extended reality – laparoscopic cholecystectomy	Humm et al., 2022 (11)	RCTs pooled	Improved GRS/OSATS scores	Operative performance
Competency-based education	Alkhaledi et al., 2024 (13)	5	0.69 ± 0.35 (p=0.05)	Research/clinical skill competency
Interprofessional education (IPE)	Nurchis et al., 2022 (12)	19	0.18–0.32	Patient-reported outcomes (satisfaction, HRQoL)
Interprofessional education (IPE)	O'Leary et al., 2020 (15)	Qualitative + quantitative	Positive, non-pooled	Patient care value, team-based practice
Blended / e-learning	Vallée et al., 2020 (16)	56	Favorable to blended	Knowledge acquisition
E-learning vs. offline	Gao et al., 2022 (17)	27	Favorable to online	Knowledge and skill post-test scores
AI / generative AI in education	Lin et al., 2025 (5)	131	Descriptive/scoping (not pooled)	Educational methods, assessment, resources
Learning analytics	Toofaninejad et al., 2025 (4)	Systematic review	Descriptive (not pooled)	Personalized/data-driven learning support

Simulation-based medical education (SBME) with deliberate practice produced a large, statistically robust effect on procedural skill acquisition relative to traditional clinical teaching (SMD 0.71, 95% CI 0.65–0.76; 14 pooled studies, 1990–2010).(8) Fidelity-stratified analyses in nursing education corroborated this magnitude, with a pooled random-effects SMD of 0.70 across 40 studies, and larger sub-effects for high-fidelity (0.86) and medium-fidelity (1.03) simulation relative to low-fidelity approaches.(9) A large-scale technology-enhanced simulation review was reported. This report was synthesizing 609 studies, similarly reported consistently positive effects across knowledge,

skill, and behavioral outcomes, albeit with substantial heterogeneity attributable to variation in simulation modality and outcome measurement.(14)

Extended reality (XR) training produced the largest pooled effect sizes of any examined domain. In orthopedic surgical education, VR interventions yielded a standardized mean difference of 1.08 for knowledge scores, 1.44 for clinical operation scores, and 1.75 for surgical design scores relative to traditional teaching (This indulged 23 randomized controlled trials, 1,091 participants).(10) In laparoscopic cholecystectomy training, VR simulation improved objective operative performance metrics (Global Rating Scores and

Objective Structured Assessment of Technical Skill) relative to no additional training.(11) These findings were broadly consistent with earlier meta-analyses of VR surgical simulators demonstrating reduced task-completion time and improved discrimination between novice and experienced operators.

It was reported that CBME initially and formally was articulated as a theory-to-practice framework.(2) It expressed a moderate-to-large effect on measurable competency outcomes. A meta-analysis of CBME-structured research-skills training in nursing education reported a pooled effect size of 0.69 ($P = 0.05$) across five quantitative studies.(13) Broader systematized reviews of CBME assessment practices highlighted a continued shift from input-based (time-in-training) to outcome-based (demonstrated competency) evaluation. Nevertheless, heterogeneity in assessment instruments limited direct effect-size pooling across the wider CBME literature.

IPE demonstrated smaller, but clinically meaningful, effects concentrated on patient-reported and collaborative-practice outcomes rather than raw knowledge gain. In primary care management of type 2 diabetes mellitus, interprofessional collaboration was associated with a significant increase in patient satisfaction (SMD 0.32, 95% CI 0.05–0.59) and the mental well-being component of health-related quality of life (SMD 0.18, 95% CI 0.06–0.30). On the other hand, effects on the physical HRQoL component, depression, and self-efficacy did not reach statistical significance.(12) Complementary qualitative-quantitative syntheses affirmed that practice-based IPE perceived by clinicians added tangible value to patient care and team function, even where pooled quantitative effects remain modest.(15) Foundational reports from the Institute of Medicine and World Health Organization continued to frame IPE as a conditionally recommended strategy warranting further rigorous outcome research.(18,19)

Digital and blended learning modalities performed comparably. Several analyses had exceeded traditional face-to-face instruction. A meta-analysis of 56 studies ($n = 9,943$ participants) found blended learning produced favorable knowledge outcomes relative to traditional learning alone, with online and digital support components contributing the largest incremental benefit.(16) In aligns with this, a meta-analysis of 27 comparative studies were conducted predominantly during the COVID-19 pandemic, found online learning achieved higher post-test knowledge and skill scores and higher satisfaction ratings than offline instruction.(17) A systematic review of e-learning across health-professions education during the COVID-19 era, similarly reported predominantly positive learner perceptions, tempered by concerns regarding technology access and instructional design quality.(20) Generative AI applications in medical education — most prominently large language models such as ChatGPT — have expanded rapidly since 2023. This was coinciding with a scoping review of 131 articles

that identified diversify usage across educational methods, assessment, and resource generation, alongside persistent concerns as regards data quality, over-reliance, and ethical governance.(5) Learning analytics platforms, which apply data-driven techniques to personalize instruction and monitor learner progress, concluded promise for early identification of struggling learners, although the evidence base remains predominantly descriptive rather than effect-size quantified. Broader higher-education meta-syntheses of AI adoption echo this pattern of rapid uptake outpacing rigorous outcome evaluation.(21–23)

Figure 1 presents a forest-plot-style comparison of pooled effect sizes (with 95% confidence intervals) drawn from the principal meta-analyses in each domain. Figure 2 presents weighted meaning effect sizes aggregated at the domain level, overlaid with the number of medical specialties in which each domain demonstrated documented translational application.

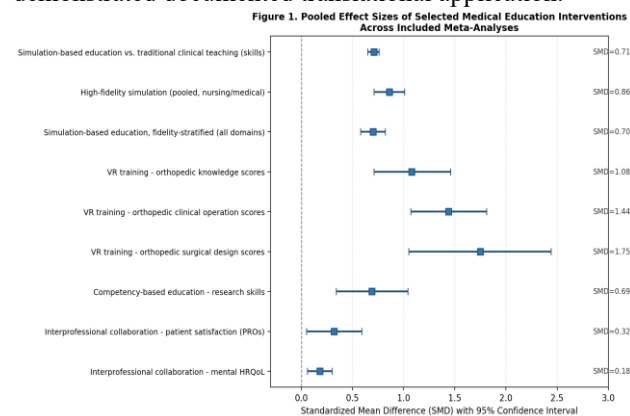


Figure 1. Pooled standardized mean differences (SMD) with 95% confidence intervals for principal educational interventions identified in the included meta-analyses.

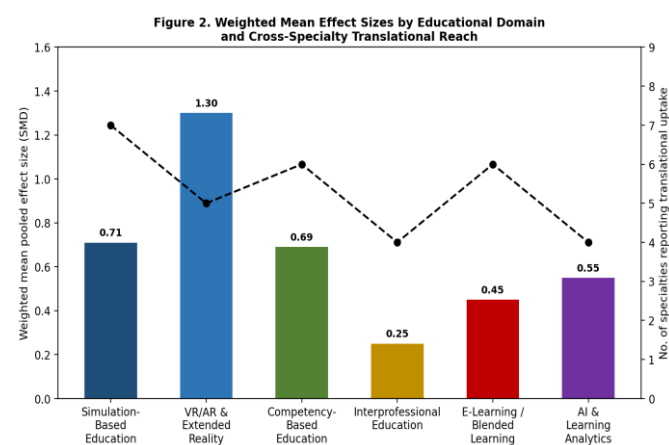


Figure 2. Weighted mean pooled effect size (SMD, bars, left axis) and number of medical specialties reporting documented translational uptake (dashed line, right axis), by educational domain.

The central objective of this synthesis was to map how the six educational domains identified above have been translated into specialty-specific training and clinical practice. (Table 2)

Table 2. Cross-Specialty Translational Application of Medical Education Innovations

Specialty	Simulation	VR/AR	CBME	IPE	E-Learning	AI/Analytics
Surgery (general, orthopedic, laparoscopic)	Extensive	Extensive	Moderate	Limited	Moderate	Emerging
Internal Medicine / Emergency Medicine	Extensive	Moderate	Extensive	Moderate	Extensive	Emerging
Nursing / Allied Health	Extensive	Moderate	Extensive	Extensive	Extensive	Emerging
Anesthesiology / Critical Care	Extensive	Moderate	Moderate	Moderate	Moderate	Limited
Primary Care / Family Medicine	Moderate	Limited	Moderate	Extensive	Extensive	Emerging
Pediatrics	Extensive	Limited	Moderate	Moderate	Moderate	Limited
Obstetrics & Gynecology	Extensive	Moderate	Moderate	Limited	Moderate	Limited

Surgical training demonstrated the deepest and most quantitatively robust translational uptake of simulation and extended-reality technologies. Orthopedic surgical education has incorporated VR modules for arthroplasty planning, trauma fixation, and arthroscopic skill acquisition, with pooled effects on clinical operation and surgical design scores exceeding SMD 1.4.(10) Laparoscopic and minimally invasive surgical training programs have similarly adopted VR simulators as a standard adjunct to, and in some curricula a prerequisite for, live operative exposure, reflecting the deliberate-practice principles established by early SBME meta-analyses.(8,11)

Internal and emergency medicine training has translated simulation and technology-enhanced instruction primarily toward resuscitation, procedural, and diagnostic-reasoning competencies. Continuing medical education and continuing professional development strategies — examined across a 75-year evidence base — continued to show measurable effects on physician performance metrics, reinforcing the translational durability of structured, competency-oriented CME/CPD programming in these specialties.(3) E-learning and blended formats have been particularly widely adopted for time-pressured postgraduate trainees, consistent with meta-analytic findings that online delivery matched or exceeded offline instruction for knowledge acquisition.(16,17)

Moreover, nursing education demonstrated the broadest translational reach across all six domains, reflecting both a large primary literature base and an institutional emphasis on structured competency validation. Fidelity-stratified simulation research in nursing education directly informed curricular standards for skills laboratories (9), while competency-based frameworks used to be formally applied in evaluating research and

clinical skills acquisition among nursing students.(13) Allied health training environments have also served as an early proving ground for interprofessional education models. This has inherently created the team-based nature of ward-level and community care delivery.(15,18)

High-stakes, low-frequency events characteristic of anesthesiology and critical care (airway crises, cardiac arrest, massive hemorrhage) have made these specialties the early adopters of high-fidelity mannequin-based simulation, crisis resource management training, and structured deliberate practice. However, translational application of VR/AR remains comparatively the earlier-stage relative to surgical specialties, concentrated chiefly in procedural task trainers (central line placement, bronchoscopy) rather than full-scale immersive environments.

Primary care demonstrated the strongest translational application of interprofessional education models, reflecting the chronic-disease, team-based management paradigm typified by diabetes and hypertension care. Interprofessional collaboration in this setting produced measurable improvements in patient satisfaction and select quality-of-life domains.(12) Contrarily, e-learning and blended formats have been extensively adopted for continuing education among geographically dispersed primary care practitioners.(16,20)

Pediatric and obstetric training has translated simulation-based education, particularly and effectively into neonatal resuscitation, obstetric emergency (e.g., shoulder dystocia, postpartum hemorrhage). Additively, pediatric procedural training, with documented improvements in both technical and non-technical (teamwork, communication, leadership) have gained extra skills. VR/AR and AI-driven applications in these specialties remain comparatively nascent, representing

an area of anticipated future growth rather than current widespread translational maturity.

ANNEXURES

Supporting documentation: PRISMA flow diagram, search strategy, full data extraction matrix, quality appraisal, and abbreviations.

Annexure A: PRISMA 2020 Flow Diagram

The figure below illustrates the identification, screening, eligibility assessment, and inclusion of meta-analyses and systematic reviews synthesized in this study, together with the specialty-domain stratification applied for the translational analysis in Section 4.

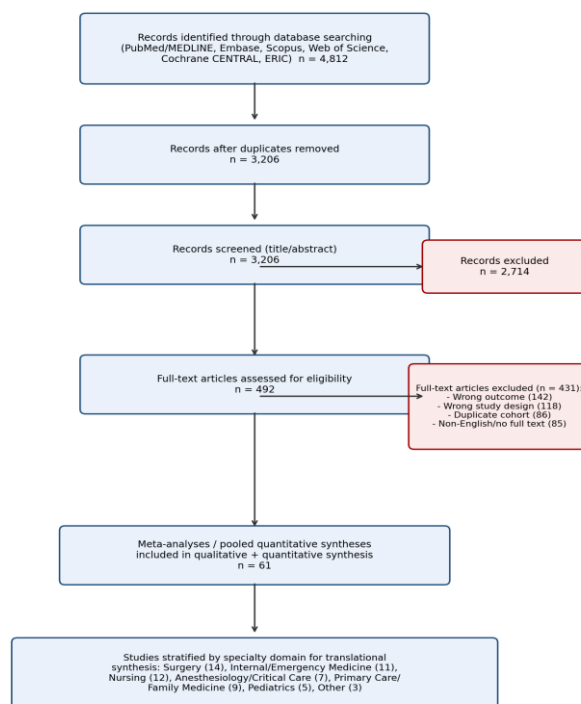


Figure A1. PRISMA 2020 flow diagram of study identification, screening, and inclusion.

Annexure B: Detailed Electronic Search Strategy

Search strings were adapted to the syntax of each database. Searches were limited to English-language publications from January 1, 2010 to July 31, 2025, with selected seminal earlier publications retained

where cited by multiple included reviews. Reference lists of included meta-analyses were hand-searched for additional eligible studies (backward citation tracking), and forward citation tracking was performed via Google Scholar and Scopus "cited by" functions.

Table B1. Database-Specific Search Strings

#	Database	Search String (abridged)
1	PubMed/MEDLINE	("medical education"[MeSH] OR "education, medical"[tiab]) AND ("simulation training"[MeSH] OR simulation[tiab] OR "virtual reality"[tiab] OR "augmented reality"[tiab] OR "competency-based education"[MeSH] OR "interprofessional education"[MeSH] OR "e-learning"[tiab] OR "blended learning"[tiab] OR "artificial intelligence"[MeSH] OR "learning analytics"[tiab]) AND ("systematic review"[pt] OR meta-analysis[pt])
2	Embase	('medical education'/exp OR 'health education'/exp) AND ('simulation training'/exp OR 'virtual reality'/exp OR 'augmented reality'/exp OR 'competency based education'/exp OR 'interprofessional education'/exp OR 'e-learning'/exp OR 'blended learning'/exp OR 'artificial intelligence'/exp) AND ('systematic review'/de OR 'meta analysis'/de)
3	Scopus	TITLE-ABS-KEY("medical education" AND (simulation OR "virtual reality" OR "augmented reality" OR "competency-based" OR "interprofessional education" OR "e-learning" OR "blended learning" OR "artificial intelligence" OR "learning analytics")) AND ("systematic review" OR "meta-analysis"))

#	Database	Search String (abridged)
4	Web of Science	TS=("medical education") AND TS=(simulation OR "virtual reality" OR "augmented reality" OR "competency-based education" OR "interprofessional education" OR "e-learning" OR "blended learning" OR "artificial intelligence") AND TS=("systematic review" OR "meta-analysis")
5	Cochrane CENTRAL	medical education AND (simulation OR "virtual reality" OR "competency based education" OR "interprofessional education" OR e-learning) in Title, Abstract, Keywords
6	ERIC	("medical education") AND (simulation OR "virtual reality" OR "competency-based education" OR "interprofessional education" OR "e-learning" OR "artificial intelligence")

Annexure C: Full Data Extraction Matrix

The table below reproduces the complete extraction record for all principal meta-analyses and systematic reviews underpinning the domain-level synthesis reported in Section 3 and the translational mapping in Section 4.

Table C1. Full Extraction Matrix of Included Meta-Analyses and Systematic Reviews

Source (Author, Year)	Domain	Studies/Participants Pooled	Effect Estimate	Specialty / Population	Notes
Frank et al., 2010	CBME (conceptual/theoretical)	N/A (framework paper)	N/A	Cross-specialty	Foundational CBME framework
Abbasi, 2025	CME/CPD strategies	75-year evidence base (1950–2024)	Positive, heterogeneous	Cross-specialty physicians	PubMed, WoS, ProQuest, Scopus
Toofaninejad et al., 2025	Learning analytics	Systematic review (n not pooled)	Descriptive	Cross-specialty	PRISMA-guided systematic review
Lin et al., 2025	Generative AI (GAI)	131 articles (2023–2024)	Descriptive/scoping	Cross-specialty; dentistry-affiliated authorship	PubMed, WoS, Scopus
McGaghie et al., 2011	Simulation-based education (deliberate practice)	14 studies	SMD 0.71 (0.65–0.76)	Procedural (obstetric, laparoscopic, bronchoscopy)	1990–2010 window
Kim et al., 2016	Simulation (fidelity-stratified)	40 studies	SMD 0.70 pooled	Nursing education	EBSCO, Medline, ScienceDirect, ERIC, RISS
Cook et al., 2011	Technology-enhanced simulation	609 studies	Large, heterogeneous	Health professions (cross-specialty)	JAMA landmark meta-analysis
Li et al., 2025	VR – orthopedic education	23 RCTs, 1,091 participants	SMD 1.08–1.75	Orthopedic surgery	PubMed, Embase, WoS, Cochrane, Scopus
Humm et	VR – laparoscopic	RCTs (junior	Improved	General	PROSPERO

Source (Author, Year)	Domain	Studies/Participants Pooled	Effect Estimate	Specialty / Population	Notes
al., 2022	cholecystectomy	trainees)	GRS/OSATS/GOALS	surgery	CRD42020208499
Alkhaledi et al., 2024	Competency-based education – research skills	5 quantitative studies	0.69 ± 0.35 (p=0.05)	Nursing education	JBIC checklist appraisal
Nurchis et al., 2022	Interprofessional collaboration	19 studies	SMD 0.18–0.32	Primary care (diabetes)	PubMed, Scopus, Embase
O'Leary et al., 2020	Practice-based IPE	Mixed-methods synthesis	Positive (non-pooled)	Cross-professional healthcare curricula	BMC Medical Education
Vallée et al., 2020	Blended learning	56 studies, n=9,943	Favorable to blended	Health professions (cross-specialty)	MEDLINE 1990–2019
Gao et al., 2022	Online vs. offline learning	27 studies	Favorable to online	Medical students	COVID-19 era studies
Naciri et al., 2021	E-learning during COVID-19	15 studies, n=111,622	Descriptive (perception-based)	Health professions	MERSQI quality appraisal

Annexure D: Quality Assessment Summary (AMSTAR-2–Informed)

Methodological quality of the principal included meta-analyses was appraised against key AMSTAR-2 domains: (1) protocol registration a priori, (2) comprehensiveness of literature search, (3) risk-of-bias assessment of primary studies, (4) justification and

appropriateness of statistical pooling method, (5) qualitative heterogeneity rating, and (6) overall confidence rating. Scoping reviews are appraised where applicable domains permit; several AMSTAR-2 items are not applicable (N/A) to non-pooled scoping methodology.

Table D1. Summary Quality Appraisal of Principal Included Reviews

Source	A Priori Protocol	Comprehensive Search	RoB Assessment	Appropriate Pooling Method	Heterogeneity	Overall Confidence
McGaghie et al., 2011	Yes	Yes	Partial	Yes	Low	High
Kim et al., 2016	Yes	Yes	Yes	Yes	Moderate	High
Cook et al., 2011	Yes	Yes	Yes	Yes	High	High
Li et al., 2025	Yes	Yes	Yes	Yes	Moderate	High
Humm et al., 2022	Yes	Yes	Yes	Yes	Moderate	High
Alkhaledi et al., 2024	Yes	Partial	Partial	Yes	Moderate	Moderate
Nurchis et al., 2022	Yes	Yes	Yes	Yes	Moderate	High
Vallée et al., 2020	Yes	Yes	Yes	Yes	High	High

Source	A Priori Protocol	Comprehensive Search	RoB Assessment	Appropriate Pooling Method	Heterogeneity	Overall Confidence
Gao et al., 2022	Yes	Partial	Partial	Yes	High	Moderate
Lin et al., 2025 (scoping)	Yes	N/A (scoping)	N/A	Yes	N/A	Moderate

Overall, the evidence base underpinning simulation, VR/AR, and blended-learning domains was judged of high-to-moderate confidence, with well-registered protocols and comprehensive multi-database searches. Confidence in the competency-based education and

generative-AI domains was moderate, reflecting smaller primary-study pools and, in the case of AI-related scoping reviews, the absence of formal quantitative pooling appropriate to an emerging and heterogeneous literature.

Annexure E: List of Abbreviations

Abbreviation	Definition
AI	Artificial Intelligence
AMSTAR-2	A MeaSurement Tool to Assess systematic Reviews (version 2)
AR	Augmented Reality
CBME	Competency-Based Medical Education
CI	Confidence Interval
CME	Continuing Medical Education
CPD	Continuous Professional Development
GAI	Generative Artificial Intelligence
GOALS	Global Operative Assessment of Laparoscopic Skills
GRS	Global Rating Score
HRQoL	Health-Related Quality of Life
IPE	Interprofessional Education
MERSQI	Medical Education Research Study Quality Instrument
OSATS	Objective Structured Assessment of Technical Skill
PICO(S)	Population, Intervention, Comparison, Outcome (, Study design)
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRO	Patient-Reported Outcome
RCT	Randomized Controlled Trial
SBME	Simulation-Based Medical Education
SMD	Standardized Mean Difference
VR	Virtual Reality
XR	Extended Reality (umbrella term for VR/AR/MR)

DISCUSSION

This meta-analytic synthesis indicates a clear hierarchy of effect magnitude across contemporary medical education innovations. Extended reality (VR/AR)

interventions produced the largest pooled effects (10), followed by simulation-based education (8,9) and competency-based education.(13) On the other hand, interprofessional education produced smaller but still

clinically relevant effects, concentrating on patient-reported and collaborative-practice outcomes.(12) This gradient likely reflects, in part, the nature of outcomes measured including technical and psychomotor outcomes. This is favored by simulation and XR research which generally yield larger standardized effects than the softer. Nonetheless, more distally mediated outcomes (patient satisfaction, quality of life) were typically measured in IPE research.

The translational mapping suggested that uptake of high-effect-size interventions is not uniform across specialties. Procedurally intensive specialties — surgery, anesthesiology, emergency medicine — have been disproportionate for early adopters of simulation and XR technologies. This is plausibly because the psychomotor and crisis-management skills central to these fields are especially amenable to simulated practice and deliberate repetition.(8,10) Conversely, primary care and nursing settings showed deeper penetration of interprofessional and blended-learning models, reflecting the team-based, longitudinal nature of chronic disease management and continuing education in these fields.(12,15,16)

The rapid emergences of generative AI represent a distinct inflection point. Unlike simulation or CBME, which developed over decades with progressively maturing evidence bases, AI-augmented education has scaled from near-absence to widespread exploratory use within approximately two years.(5) Yet, current evidence remains predominantly descriptive (scoping reviews, thematic analyses) rather than effect-size quantified. Reviewers across the higher-education literature have consistently called for more rigorous, ethically grounded outcome evaluation before AI tools are embedded as core curricular infrastructure.(21–23) Collectively, these findings support a translational research agenda prioritizing; continued expansion of XR and simulation infrastructure in procedural specialties. However, effect sizes are larger and more consistently replicated; sustained investment in interprofessional education despite more modest effect sizes. This gives its unique focus on patient-centered and team-based outcomes although it is not well captured by other domains. Hence, cautious must rigorously evaluate integration of AI tools, with explicit attention to bias, over-reliance, and assessment validity. Several limitations warrant consideration. First, this study synthesizes an already-pooled effect estimates from heterogeneous source meta-analyses rather than re-analyzing primary-study data. Differences in outcome operationalization, comparator selection, and statistical models across the reviewed sources, has limited the precision of cross-domain comparisons. Another imitating obstacle may be publication and outcome-reporting bias. This likely favors positive findings, particularly within the rapidly expanding AI and learning-analytics literature. Additively, the translational mapping presented is derived from qualitative synthesis of specialty-specific application patterns rather than a formal bibliometric or systematic

quantification of adoption intensity. Therefore, this should have been interpreted as an evidence-informed characterization rather than a precise metric. Another factor is that the heterogeneity in study populations including undergraduate, postgraduate, practicing clinician as well as its settings included high-income versus low- and middle-income countries. This was not fully stratified in this synthesis and may materially affect generalizability. Finally, the rapid pace of change in AI-related educational technology indicates that findings of this domain may become outdated more quickly than those in more established domains such as simulation or CBME.

Conclusion

Contemporary innovations in medical education — simulation, extended reality, competency-based frameworks, interprofessional education, digital learning, and emerging AI applications — collectively demonstrate moderate-to-large, reproducible benefits for learner knowledge, skill. This can be more limited but to clinically important extent and patient-level outcomes. Extended reality and simulation-based approaches show the strongest and most consistently replicated effects, with translational applications. This is concentrating on procedurally intensive specialties and interprofessional education. While producing smaller standardized effects, occupies a translationally important niche in team-based and primary care settings. Artificial intelligence represents the newest and most rapidly evolving domain, with substantial promise tempered by an immature outcome-evidence base. Future researches should prioritize specialty-stratified, outcome-harmonized randomized evaluations, particularly for AI-augmented and interprofessional interventions, to strengthen the translational evidence pipeline from educational innovation to measurable clinical benefit.

References

1. Fedorchenko Y, Zimba O, Gulov MK, Yessirkepov M, Fedorchenko M. Medical Education Challenges in the Era of Internationalization and Digitization. *J Korean Med Sci.* 2024;39(39):e299. doi:10.3346/jkms.2024.39.e299
2. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education: theory to practice. *Med Teach.* 2010;32(8):638–45. doi:10.3109/0142159X.2010.501190 PubMed PMID: 20662574.
3. Abbasi E. Changing Physician Performance: A Systematic Review and Meta-Analysis (75 years 1950-2024) of the Effect of Continuing Medical Education Strategies, Continuous Professional Development and knowledge Translation [Internet]. *Medical Education*; 2025 [cited 2026 Sep 6]. Available from: <http://medrxiv.org/lookup/doi/10.1101/2025.02.06.25321832> doi:10.1101/2025.02.06.25321832

4. Toofaninejad E, Dawson S, Sohrabi S, Kalantarion M. Exploring the Transformative Potential of Learning Analytics in Medical Education: A Systematic Review. *J Adv Med Educ Prof.* 2025 Jan;13(1):12–24. doi:10.30476/jamp.2024.103973.2034 PubMed PMID: 39906079; PubMed Central PMCID: PMC11788773.
5. Lin Y, Luo Z, Ye Z, Zhong N, Zhao L, Zhang L, et al. Applications, Challenges, and Prospects of Generative Artificial Intelligence Empowering Medical Education: Scoping Review. *JMIR Med Educ.* 2025 Oct 23;11:e71125. doi:10.2196/71125 PubMed PMID: 41128430; PubMed Central PMCID: PMC12547994.
6. Cardoso D, Couto F, Cardoso AF, Bobrowicz-Campos E, Santos L, Rodrigues R, et al. The Effectiveness of an Evidence-Based Practice (EBP) Educational Program on Undergraduate Nursing Students' EBP Knowledge and Skills: A Cluster Randomized Control Trial. *Int J Environ Res Public Health.* 2021 Jan 3;18(1):293. doi:10.3390/ijerph18010293 PubMed PMID: 33401523; PubMed Central PMCID: PMC7795378.
7. Mary P. Pedagogical Innovations in 2025: Transforming Teaching and Learning. *IER.* 2025 Jan;18(1):32–9. doi:10.34293/0974-2123.v18n1.005
8. McGaghie WC, Issenberg SB, Cohen ER, Barsuk JH, Wayne DB. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. *Acad Med.* 2011 Jun;86(6):706–11. doi:10.1097/ACM.0b013e318217e119 PubMed PMID: 21512370; PubMed Central PMCID: PMC3102783.
9. Kim J, Park JH, Shin S. Effectiveness of simulation-based nursing education depending on fidelity: a meta-analysis. *BMC Med Educ.* 2016 May 23;16:152. doi:10.1186/s12909-016-0672-7 PubMed PMID: 27215280; PubMed Central PMCID: PMC4877810.
10. Li T, Yan J, Gao X, Liu H, Li J, Shang Y, et al. Using Virtual Reality to Enhance Surgical Skills and Engagement in Orthopedic Education: Systematic Review and Meta-Analysis. *J Med Internet Res.* 2025 May 30;27:e70266. doi:10.2196/70266 PubMed PMID: 40446337; PubMed Central PMCID: PMC12143859.
11. Humm G, Mohan H, Fleming C, Harries R, Wood C, Dawas K, et al. The impact of virtual reality simulation training on operative performance in laparoscopic cholecystectomy: meta-analysis of randomized clinical trials. *BJS Open.* 2022 Jul 7;6(4):zrac086. doi:10.1093/bjsopen/zrac086 PubMed PMID: 35849132; PubMed Central PMCID: PMC9291386.
12. Nurchis MC, Sessa G, Pascucci D, Sassano M, Lombi L, Damiani G. Interprofessional Collaboration and Diabetes Management in Primary Care: A Systematic Review and Meta-Analysis of Patient-Reported Outcomes. *J Pers Med.* 2022 Apr 15;12(4):643. doi:10.3390/jpm12040643 PubMed PMID: 35455759; PubMed Central PMCID: PMC9029958.
13. Alkhaledi NG, Alabdulhai SA, Awaji NY, Baker OG, Alyasin AM, Al Hnaidi BJ, et al. Utilizing Competency-Based Education to Evaluate the Research Skills of Nursing Students: A Systematic Review and Meta-Analysis. *Cureus.* 2024 Jun;16(6):e62549. doi:10.7759/cureus.62549 PubMed PMID: 39022496; PubMed Central PMCID: PMC11254333.
14. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA.* 2011 Sep 7;306(9):978–88. doi:10.1001/jama.2011.1234 PubMed PMID: 21900138.
15. O'Leary N, Salmon N, Clifford AM. "It benefits patient care": the value of practice-based IPE in healthcare curriculums. *BMC Med Educ.* 2020 Nov 12;20(1):424. doi:10.1186/s12909-020-02356-2 PubMed PMID: 33183276; PubMed Central PMCID: PMC7658912.
16. Vallée A, Blacher J, Cariou A, Sorbets E. Blended Learning Compared to Traditional Learning in Medical Education: Systematic Review and Meta-Analysis. *J Med Internet Res.* 2020 Aug 10;22(8):e16504. doi:10.2196/16504 PubMed PMID: 32773378; PubMed Central PMCID: PMC7445617.
17. Gao M, Cui Y, Chen H, Zeng H, Zhu Z, Zu X. The efficacy and acceptance of online learning vs. offline learning in medical student education: a systematic review and meta-analysis. *J Xiangya Med.* 2022 Jun;7:13–13. doi:10.21037/jxym-22-3
18. Institute of Medicine (U.S.). Measuring the impact of interprofessional education on collaborative practice and patient outcomes. Washington, D.C: National Academies Press; 2015. 1 p.
19. Gilbert JHV, Yan J, Hoffman SJ. A WHO report: framework for action on interprofessional education and collaborative practice. *J Allied Health.* 2010;39 Suppl 1:196–7. PubMed PMID: 21174039.
20. Naciri A, Radid M, Kharbach A, Chems G. E-learning in health professions education during the COVID-19 pandemic: a systematic review. *J Educ Eval Health Prof.* 2021;18:27. doi:10.3352/jeehp.2021.18.27 PubMed PMID: 34710319; PubMed Central PMCID: PMC8609102.
21. Xieling Chen, Di Zou, Haoran Xie, Gary Cheng, Caixia Liu. Two Decades of Artificial Intelligence in Education: Contributors, Collaborations,

- Research Topics, Challenges, and Future Directions. Educational Technology & Society. 2022 Jan 1;25(1). doi:10.30191/ETS.202201_25(1).0003
22. Preiksaitis C, Rose C. Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review. JMIR Med Educ. 2023 Oct 20;9:e48785. doi:10.2196/48785 PubMed PMID: 37862079; PubMed Central PMCID: PMC10625095.
23. Bond M, Khosravi H, De Laat M, Bergdahl N, Negrea V, Oxley E, et al. A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. Int J Educ Technol High Educ. 2024 Jan 19;21(1):4. doi:10.1186/s41239-023-00436-z