Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Procedural fluency and technical performance in cardiopulmonary resuscitation training: A piecewise longitudinal analysis

Hien Nguyen Trong
Pham Ngoc Thach University of Medicine
Anh Nguyen Thi Lan
Pham Ngoc Thach University of Medicine
Quynh Nguyen Nhu
Pham Ngoc Thach University of Medicine
Linh Nguyen Thi My
Pham Ngoc Thach University of Medicine
Phuong Bui Phan Quynh
Pham Ngoc Thach University of Medicine
Thang Hoang Quoc
Pham Ngoc Thach University of Medicine
Thu Nguyen Ngoc Phuong
nguyenngocphuongthu@pnt.edu.vn

Pham Ngoc Thach University of Medicine

Research Article

Keywords: Cardiopulmonary resuscitation, Simulation training, Medical education, Nursing education, Assessment

Posted Date: May 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9466042/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose

This study compared longitudinal cardiopulmonary resuscitation (CPR) skill trajectories between first-year medical and nursing students and examined performance changes associated with a final high-stakes assessment.

Methods

We conducted a four-time-point repeated-measures study with 211 first-year students (103 nursing, 108 medical). Performance was assessed immediately post-training (W1), week 1 (W2), week 2 (W3), and week 4 (W4, final assessment). Outcomes included a 21-item observational checklist (CHK) and device-derived quality CPR (QCPR) metrics. Piecewise generalized estimating equation models were applied to compare Phase 1 (W1–W3) and Phase 2 (W3–W4), adjusting for sex and body mass index.

Results

Both groups improved significantly over time but demonstrated distinct trajectories in procedural competence and compression rate consistency ($p_{int} < 0.001$). Medical students showed rapid early gains during Phase 1 (3.44 CHK points/step), plateauing by W3. Nursing students progressed more gradually in Phase 1 (1.69 points/step) but accelerated sharply during Phase 2 (3.52 points/step; ($p_{int} < 0.001$), coinciding with the final exam. In contrast, core technical metrics showed convergent trajectories across groups. Although all students met checklist criteria by W4, only 68.9% of nursing and 74.1% of medical students achieved adequate compression depth, indicating a discrepancy between procedural fluency and technical short-term performance optimization. Correlations between CHK and QCPR metrics were weak to moderate, underscoring the need for multimodal assessment.

Conclusion

Assessment timing was associated with discipline-specific performance changes, particularly in procedural skills. However, persistent deficits in compression depth highlight discrepancies between checklist-based performance and technical quality, supporting the integration of objective metrics in competency-based CPR training.

INTRODUCTION

Cardiac arrest is a time-critical emergency in which the quality of cardiopulmonary resuscitation (CPR) directly determines perfusion to vital organs and overall survival outcomes [1]. High-quality CPR (QCPR)

requires timely recognition, a correct sequence of basic life support actions, and chest compressions delivered within strict guideline targets for rate, depth, and recoil [2]. However, resuscitation skills are known to deteriorate rapidly without sufficient reinforcement, supporting the use of distributed practice and planned refreshers to maintain competency [3–5].

Clinical simulation is ideal for CPR training as it enables deliberate practice under controlled conditions [6]. A significant advancement in this field is the integration of objective feedback devices (QCPR), which provide real-time quantitative data on compression mechanics [7]. While such feedback is known to improve immediate performance, evidence regarding skill retention across multiple follow-up points remains limited, particularly concerning how learning trajectories may differ between diverse groups, such as medical and nursing students [8–10].

Beyond simple retention, modern educational theory suggests that assessment time points themselves, especially final examinations, can function as active learning events through the testing or retrieval effect [11, 12]. This phenomenon can produce an assessment-induced "reinforcement effect", where the act of being tested reinforces skills more effectively than natural retention alone [13]. Identifying these phase-specific patterns requires sophisticated longitudinal designs and phase-based modeling (piecewise) to separate assessment-related reinforcement from underlying skill decay.

Furthermore, selecting appropriate outcomes is critical for interpreting CPR competence. While instructor-rated checklists (CHK) capture procedural sequence and decision-making, they often exhibit ceiling effects and may overestimate actual technical performance compared to device-derived metrics [14]. Conversely, QCPR metrics directly quantify compression quality with high objectivity. Therefore, a multi-source approach combining both modalities is essential for a comprehensive understanding of a learner's practical performance.

To address these gaps, we conducted a four-time-point repeated-measures study to characterize the distinct learning trajectories of first-year nursing and medical students following simulation-based training. This research specifically aims to: (1) describe and compare longitudinal trajectories of CHK scores and objective QCPR metrics; (2) apply piecewise generalized estimating equation (GEE) modeling to isolate performance changes associated with the final high-stakes assessment; and (3) identify persistent gaps between procedural fluency and technical short-term performance optimization to inform time-targeted, discipline-specific reinforcement strategies in resuscitation education.

METHODS

1. Study design and setting

We conducted a prospective, four-time-point repeated-measures study at a Clinical Simulation Center of the university. Student performance was tracked longitudinally at four intervals: W1 (immediately post-training), W2 (first week), W3 (second week), and W4 (fourth week - final assessment). This design was specifically chosen to characterize longitudinal learning trajectories and evaluate the performance

changes during the assessment phase of the final assessment. The center previously held Society for Simulation in Healthcare Accreditation for Core, Education/ Teaching, Assessment and Research.

Clinical trial number: not applicable.

2. Participants

A total of 211 undergraduate students (103 Nursing; 108 Medical) were included. Participants were eligible if they completed the full standardized curriculum and all four assessment time points, with unique identifiers to enable individual-level data linkage.

3. Simulation training and assessment procedures

The curriculum followed a standardized workflow: a 90-minute CPR lecture and 30-minute questions and answers, followed by 270 minutes of hands-on deliberate practice on a QCPR-enabled manikin with coaching from instructors and technicians, and a final 30-minute guided debrief. During assessments (W1–W4), real-time manikin feedback was blinded to ensure the results reflected independent learner competence. Each student performed a standardized 2-minute 30:2 adult basic life support scenario, which included chest compressions, rescue breaths, and automated external defibrillator actions prompted by the device. Faculty raters underwent calibration sessions to ensure scoring consistency using standardized checklist anchors.

4. Measurement and outcome variables

We utilized a multi-source approach to capture both procedural sequence and compression mechanics:

- Observational outcome: A 21-item checklist score (range 0–42). Passing was set at CHK ≥ 24 based on a 55.9% Angoff method-derived minimum standard.
- Device-derived metrics (QCPR): Six quantitative metrics were extracted (mean compression rate (MCR, cpm); compression rate-in-range (CCR, %); mean compression depth (MCD, mm); compression depth-in-range (CCD, %); full release (FR, %); and chest compression fraction (CCF, %). Guideline-based targets included MCR (100–120 cpm), MCD (50–60 mm), and CCF (≥ 60%).

5. Statistical analysis and modeling of trajectories

To map the learning trajectories of our participants while accounting for the repeated-measures nature of the data, we used GEE. This approach is well-suited for longitudinal studies because it properly handles the correlation between multiple measurements taken from the same student over time [15, 16]. We specified an autoregressive working correlation structure, which reflects the reasonable assumption

that assessments closer together in time, such as W1 and W2, are more strongly related than those farther apart.

Our primary models compared CPR skill development between nursing and medical students across the four time points, including a Group x Time interaction term to test whether their learning trajectories differed significantly. To ensure a fair comparison, we adjusted all models for baseline physical characteristics known to influence CPR performance—specifically sex and body mass index (BMI).

We applied Gaussian GEE models for continuous outcomes, such as checklist scores and QCPR metrics (compression depth, rate, and recoil...), and logistic GEE models for binary outcomes, such as pass/fail status against guideline-based targets.

A key analytical goal was to quantify performance changes during Phase 2 (W3–W4), which coincided with preparation for and participation in the final high-stakes assessment. To capture this effect, we used piecewise (or segmented) modeling with a "knot" placed at W3. W3 was selected as the knot to represent the transition from natural skill retention to the active preparation phase for the final evaluation. This allowed us to split the study period into two distinct phases:

- Phase 1 (W1–W3): Reflecting natural skill acquisition and decay in the absence of high-stakes assessment.
- Phase 2 (W3–W4): Capturing the immediate change in performance associated with preparation for and participation in the final evaluation.

Because we examined multiple QCPR metrics simultaneously, we controlled for the increased risk of false-positive findings by applying the Benjamini-Hochberg procedure to adjust p-values and limit the false discovery rate. All data processing and statistical analyses were conducted in Python (version 3.12) using the pandas, numpy, and statsmodels libraries.

RESULTS

The final analytic sample comprised 211 first-year students (103 nursing, 108 medical) who completed all four weekly assessments and provided a complete longitudinal dataset. Baseline characteristics were largely comparable between groups in terms of age and BMI; however, the medical cohort included a higher proportion of male participants and had significantly greater average height and weight (Table 1).

Table 1
Participant Characteristics by Group

Characteristic	Nursing (n = 103)	Medical (n = 108)
Age, y	17.9 (1.8)	17.9 (1.8)
Male, n (%)	28 (27.2)	63 (58.3)
Height, cm	162.0 (7.2)	165.8 (8.9)
Weight, kg	57.9 (12.5)	62.7 (14.2)
BMI, kg/m^2	22.0 (4.0)	22.6 (3.6)
Data are presented as mean (standard deviation) unless otherwise indicated.		
BMI: body mass index.		

Both groups demonstrated substantial improvement in checklist performance over the four-week curriculum (Table 2). In adjusted GEE models (Table 3), mean CHK scores increased significantly at each follow-up relative to W1. However, their learning trajectories diverged markedly ($p_{int} < 0.001$). Medical students exhibited rapid early acquisition, gaining 6.88 points by W3 - a rate of 3.44 points per week during phase 1. In contrast, nursing students progressed more gradually in this phase (+ 3.39 points by W3, or 1.69 points/week), but then displayed a pronounced acceleration in phase 2, surging by 7.21 points - equivalent to 3.52 points per week. By W3, medical students had higher procedural scores (adjusted mean difference 3.49 points; 95% CI: 2.51, 4.47; $p < 0.001$), consistent with faster early skill acquisition. Notably, the medical group's procedural gains plateaued in Phase 2 (slope = 1.08, $p = 0.003$), while nursing students experienced a sharp upturn. This reversal (pint < 0.001) suggests that the final high-stakes assessment was temporally associated with a more pronounced performance upturn in nursing students, likely due to their lower baseline proficiency entering phase 2.

Table 2
Attainment of Pre-specified CPR Competency Targets by Assessment Time Point and Discipline Group

Outcome (passed)	Group	W1	W2	W3	W4
CHK ≥ 24/42	Nursing	100/103 (97.1)	102/103 (99.0)	102/103 (99.0)	103/103 (100.0)
	Medical	101/108 (93.5)	106/108 (98.1)	108/108 (100.0)	108/108 (100.0)
MCR 100–120/min	Nursing	62/103 (60.2)	68/103 (66.0)	68/103 (66.0)	99/103 (96.1)
	Medical	61/108 (56.5)	83/108 (76.9)	102/108 (94.4)	105/108 (97.2)
MCD 50–60 mm	Nursing	42/103 (40.8)	64/103 (62.1)	68/103 (66.0)	71/103 (68.9)
	Medical	55/108 (50.9)	68/108 (63.0)	71/108 (65.7)	80/108 (74.1)
CCF ≥ 60%	Nursing	79/103 (76.7)	86/103 (83.5)	87/103 (84.5)	99/103 (96.1)
	Medical	78/108 (72.2)	94/108 (87.0)	95/108 (88.0)	103/108 (95.4)

Data are presented as n/N (%).

Passing criteria were prespecified: CHK ≥ 24/42 (Angoff-derived minimum standard), MCR 100–120/min, MCD 50–60 mm, and CCF ≥ 60%.

CHK: checklist score; MCR: mean compression rate (cpm); MCD: mean compression depth (mm); CCF: chest compression fraction (%); W1: immediately post-training; W2: week 1; W3: week 2; W4: week 4.

Table 3
Changes in CHK and QCPR Metrics Over Time (Gaussian GEE; adjusted for BMI and sex)

Outcome	Nursing			Medical			p_{int}
	W2–W1	W3–W1	W4–W1	W2–W1	W3–W1	W4–W1	
CHK (0–42)	3.34 (2.41, 4.27) $p < 0.001$	3.39 (2.46, 4.32) $p < 0.001$	7.21 (6.28, 8.14) $p < 0.001$	2.93 (1.88, 3.97) $p < 0.001$	6.88 (5.90, 7.86) $p < 0.001$	7.86 (6.90, 8.82) $p < 0.001$	< 0.001
CCR (%)	7.21 (-1.43, 15.86) $p = 0.102$	7.39 (-1.15, 15.93) $p = 0.090$	31.74 (25.49, 37.98) $p < 0.001$	18.43 (9.99, 26.86) $p < 0.001$	34.57 (28.25, 40.90) $p < 0.001$	38.55 (31.64, 45.45) $p < 0.001$	< 0.001
MCR (cpm)	-2.98 (-5.77, -0.19) $p = 0.036$	-3.13 (-5.94, -0.33) $p = 0.029$	-5.40 (-7.71, -3.08) $p < 0.001$	-7.42 (-10.91, -3.94) $p < 0.001$	-6.80 (-9.96, -3.64) $p < 0.001$	-7.22 (-10.18, -4.27) $p < 0.001$	0.211
CCD (%)	14.21 (6.91, 21.52) $p < 0.001$	15.22 (8.11, 22.34) $p < 0.001$	25.40 (18.17, 32.63) $p < 0.001$	9.21 (-1.17, 19.59) $p = 0.082$	19.32 (13.15, 25.50) $p < 0.001$	27.04 (20.26, 33.82) $p < 0.001$	0.206
MCD (mm)	2.50 (0.80, 4.19) $p = 0.004$	2.37 (0.91, 3.83) $p = 0.002$	4.21 (2.43, 5.99) $p < 0.001$	2.85 (0.43, 5.27) $p = 0.021$	3.82 (2.05, 5.60) $p < 0.001$	4.19 (1.91, 6.47) $p < 0.001$	0.480
FR (%)	-0.45 (-5.72, 4.83) $p = 0.868$	1.82 (-3.20, 6.83) $p = 0.478$	6.69 (1.45, 11.92) $p = 0.012$	5.57 (-2.44, 13.59) $p = 0.173$	12.27 (6.20, 18.33) $p < 0.001$	14.44 (8.40, 20.49) $p < 0.001$	0.062
CCF (%)	0.75 (-2.14, 3.63) $p = 0.611$	1.68 (-0.97, 4.33) $p = 0.214$	4.82 (2.99, 6.64) $p < 0.001$	0.12 (-3.93, 4.17) $p = 0.953$	2.22 (-0.48, 4.93) $p = 0.107$	4.92 (2.69, 7.15) $p < 0.001$	0.951

Model: Gaussian GEE with an AR(1) working correlation structure, clustered by participant. Estimates are presented as Δ versus W1 (95% CI) with p-values are FDR-adjusted (BH) across QCPR outcomes/time contrasts; pint denotes the Wald test for the Group×Time interaction (W2–W4).

Adjustment: BMI and sex; reference covariate values: BMI=22.306; sex=0.431.

CHK: checklist score; QCPR: device-derived quality cardiopulmonary resuscitation; GEE: generalized estimating equations; BMI: body mass index; W1: immediately post-training; W2: week 1; W3: week 2; W4: week 4; CCR: compression rate-in-range (%); MCR: mean compression rate (cpm); CCD: compression depth-in-range (%); MCD: mean compression depth (mm); FR: full release (%); CCF: chest compression fraction (%); AR: autoregressive; FDR: false discovery rate; BH: Benjamini–Hochberg.

Spearman correlation analyses (Table S1) revealed only weak-to-moderate associations between CHK scores and objective QCPR metrics. Critically, key technical parameters, MCR and FR, showed no statistically significant correlation with checklist performance at most time points ($p > 0.05$). For example, the correlation between FR and CHK hovered near zero ($r \approx 0.00$) during W1–W2, indicating that human observation lacks the sensitivity to detect subtle but clinically vital psychomotor deficits.

Improvements in compression mechanics followed distinct patterns across QCPR components (Table 3). Both groups successfully moderated their pace toward guideline-recommended ranges. However, medical students began improving as early as W2, while nursing students showed their largest gain only at W4 (+ 31.74 percentage points in CCR). The interaction p-value was highly significant ($p_{int} < 0.001$), confirming discipline-specific timing. MCD and CCD improved comparably in both groups (+ 4.21 mm for nursing; +4.19 mm for medical by W4). Unlike rate, depth trajectories did not differ significantly between disciplines ($p_{int} = 0.480$ for MCD), suggesting this component may be more universally trainable. FR and CCF increased over time in both cohorts. Notably, CCF gains were concentrated in Phase 2 for both groups, aligning with heightened effort during the final assessment week. Importantly, the medical group showed no significant improvement in MCD during Phase 2 (slope = 0.47, 95% CI: −1.66 to 2.60; $p = 0.666$), consistent with a ceiling effect in both procedural and technical domains.

By W4, 100% of students passed the procedural checklist, confirming a ceiling effect for algorithmic adherence (Table 2). MCR was also high (96.1% nursing; 97.2% medical). However, MCD remained the primary residual deficit: only 68.9% of nursing and 74.1% of medical students met the MCD pass standard. This underscores that technical mechanics are more resistant to short-term reinforcement than procedural sequencing, and that checklist-based pass/fail judgments may mask critical performance gaps.

All models were adjusted for sex and BMI, with p-values corrected for multiple comparisons using the Benjamini–Hochberg false discovery rate procedure (Python version 3.12).

DISCUSSION

Both groups improved significantly, but their patterns differed. Medical students gained procedural skills quickly and reached a plateau by W3. Nursing students, in contrast, showed slower initial progress but made a sharp improvement between W3 and W4 - the period that included the final high-stakes assessment. This timing suggests that preparing for the exam may have encouraged them to focus their practice and correct errors, especially since they had not yet reached the same level of baseline proficiency as their medical peers.

While our study did not assess physiological markers, some research suggests that high-stakes testing can trigger stress-related responses, such as elevated cortisol, which under certain conditions may influence memory processes [17]. However, the relationship between stress and skill performance is complex and not fully understood. In our study, the performance gain occurred specifically at W4, when the final assessment was administered, suggesting that psychological factors, such as increased motivation or targeted self-practice, likely played a more direct role than any underlying neuroendocrine mechanism.

In our experience, the “catch-up effect” and the boost from the final assessment aren’t two separate things: they actually work hand in hand. The nursing students started Phase 2 with weaker skills, which meant they had more room to improve. But what really made the difference was the pressure of that high-stakes final exam [12, 18]. It wasn’t just about passing, it pushed them to practice more intentionally, pay attention to their mistakes, and fix them quickly. This observation led us to a key insight: the same summative assessment can serve distinct functions depending on the learner’s starting point. For students who entered Phase 2 with notable skill gaps, particularly the nursing cohort, the high-stakes exam acted as a strong impetus for focused practice, creating a concentrated window for rapid improvement. In contrast, for those who had already achieved procedural fluency, such as the medical students, the exam did not drive significant learning; instead, it primarily functioned as a validation of existing competence.

When we looked beyond the checklist, we saw different patterns in technical performance. Compression depth improved similarly in both groups, suggesting this aspect of CPR can be taught effectively to all learners. However, control of compression rate diverged: medical students improved earlier, while nursing students made their largest gains only in W4. The divergence in learning trajectories suggests that not all CPR skills follow the same acquisition pattern. Compression depth, a largely force-based motor task, improved steadily with practice in both groups. In contrast, maintaining the correct compression rate, a skill requiring precise timing and self-monitoring, showed minimal gains until the final week, when the upcoming high-stakes assessment appeared to motivate deliberate practice and error correction.

A key concern emerged from the gap between what observers see and what sensors measure. By W4, all students passed the observational checklist, yet nearly 30% still failed to deliver compressions deep enough to meet current guidelines. Our findings are consistent with those reported by Lakomek et al. [19]. Indeed, nearly 30% of students who passed the observational checklist still delivered compressions too shallow to meet current guidelines, indicating that human raters can miss technically inadequate performance even when procedural steps appear correct.

After all, survival depends on compressions that are deep enough and allow full chest recoil, not just on following the correct steps [20]. These findings support combining two types of assessment: checklists to ensure safety-critical actions are performed, and objective device data to monitor the technical quality of compressions. Similarly, our findings align with those reported by Jansen [21] and Castillo [14].

In summary, our results point to three practical steps. First, once learners can reliably perform the CPR sequence, training should shift focus to objective metrics like depth and recoil, where gaps often remain hidden. Second, well-timed assessments may do more than measure learning; they can also help reinforce it. Third, because students from different disciplines start at different levels and progress at different rates, resuscitation training could be more effective if it adapts to individual or group readiness, rather than following a fixed schedule for everyone.

This study is strengthened by a longitudinal design with four repeated assessments, blinded raters to reduce observer bias, and piecewise GEE modeling that helps distinguish early learning from changes linked to the final assessment.

Several limitations should be acknowledged. First, because all participants completed a high-stakes final exam, we cannot determine whether the performance gains in W4, particularly the sharp improvement in compression rate among nursing students, were directly caused by the testing effect. Other factors, such as increased motivation before the exam or additional self-directed practice, may also have contributed. Without a control group that did not take the exam, these influences cannot be separated.

Second, the weak correlations between checklist scores and objective QCPR metrics (Table S1) show that observational ratings often miss critical technical aspects of CPR, such as adequate compression depth or full chest recoil. This suggests that checklists alone may not be sufficient for high-stakes decisions about competence.

Finally, we do not know whether the improvements seen at W4 are maintained over time. The performance peak may reflect short-term preparation for the exam rather than lasting skill acquisition. Future studies should include follow-up assessments at 3–6 months to determine whether these gains translate into sustained clinical proficiency or whether they decline after the exam.

Medical and nursing students follow distinct CPR learning trajectories: medical students achieve procedural proficiency earlier, while nursing students show delayed but accelerated gains just before a final high-stakes assessment. This timing suggests that well-placed evaluations can focus practice and support skill refinement, particularly for learners who have not yet reached baseline competence. Critically, observational checklists failed to detect persistent technical deficits (nearly 30% of students still delivered compressions too shallow by W4) highlighting the need for dual assessment that combines human observation with objective QCPR metrics. These findings support tailoring resuscitation training to learners' actual readiness rather than fixed schedules.

Declarations

Acknowledgements: We would like to extend our gratitude to Phap Tran Thai Mai, Tai Tan Cao, and Duy Anh Nguyen for their with data collection. We also acknowledge Chinh Thi Kieu Chau and Bich Thi Ngoc Cao for their administrative support and assistance with funding coordination.

Funding: This study received funding and institutional support from the university. No external funding from governmental, commercial, or nonprofit organizations was obtained.

Conflicts of interest: All authors report no financial relationships or conflicts of interest to declare.

Ethics Statement: The study was approved by the Ethics Committee in Biomedical Research at the university (No. 1441/TĐHYKPNT-HĐĐĐ).

Consent to participate: Participation was voluntary, and all participants were informed about the study purpose, confidentiality, and data protection procedures before providing written informed consent. The study was conducted in accordance with institutional ethics committee approval, and no identifying personal data were collected.

Consent to publish: Informed consent that is consent to publish was obtained from all participants.

Data Availability: De-identified participant-level data and analysis code are available from the corresponding author on reasonable request.

Author contributions: HNT: Methodology, Formal analysis, Writing – review & editing; ANTL: Project administration, Data curation, Writing – review & editing; QNN: Project administration, Data curation, Writing – review & editing; LNTM: Methodology, Writing – original draft, Writing – review & editing; PBPQ: Methodology, Writing – review & editing; THQ: Data curation, Writing – review & editing; TNNP: Funding acquisition, Formal analysis, Supervision, Writing – review & editing.

References

1. American Heart Association. Highlights of the 2020 AHA Guidelines Update for CPR and ECC. https://cpr.heart.org/-/media/cpr-files/cpr-guidelines-files/highlights/hghlghts_2020_ecc_guidelines_english.pdf. Accessed November 6, 2025.
2. European Society of Cardiology. Resuscitation in out-of-hospital cardiac arrest – it's how quickly it is done, rather than who does it. https://www.escardio.org/news/press/press-releases/resuscitation-in-out-of-hospital-cardiac-arrest-it-s-how-quickly-it-is-done-r/. Published March 2025. Accessed November 6, 2025.
3. Abebe TA, Zeleke LB, Assega MA, Sefefe WM, Gebremedhn EG. Health-Care Providers' Knowledge, Attitudes, and Practices Regarding Adult Cardiopulmonary Resuscitation at Debre Markos Referral Hospital, Gojjam, Northwest Ethiopia. Adv Med Educ Pract. 2021;12:647–654.
4. Moser DK; Coleman S. Recommendations for improving cardiopulmonary resuscitation skills retention. Heart Lung J Crit Care. 1992;21(4):372–380.
5. Chang YT; Wu KC; Yang HW, et al. Effects of different cardiopulmonary resuscitation education interventions among university students: A randomized controlled trial. PloS One. 2023;18(3):e0283099.

6. Carroll JD, Messenger JC. Medical simulation: the new tool for training and skill assessment. Perspect Biol Med. 2008;51(1):47–60.
7. Ateyyah KA, Cady CE, Poltrock JT, Pirrallo RG. A Novel Use of a Metronome in Dispatcher-assisted Cardiopulmonary Resuscitation. Prehosp Emerg Care. 2015;19(1):131-134.
8. Hattie J, Timperley H. The Power of Feedback. Review of Educational Research. 2015;77(1);81–112.
9. Anderson R, Sebaldt A, Lin Y, Cheng A. Optimal training frequency for acquisition and retention of high-quality CPR skills: A randomized trial. Resuscitation. 2019;135:153-161.
10. Silva JKB, Couto TB, Coriolano AM, Aquino A, Monte JCM. Analysis of acquisition and retention of cardiopulmonary resuscitation skills according to training frequency. Einstein (Sao Paulo). 2025;23:eAO1257.
11. Polack CW, Miller RR. Testing improves performance as well as assesses learning: A review of the testing effect with implications for models of learning. J Exp Psychol Anim Learn Cogn. 2022;48(3):222-241.
12. Larsen DP, Butler AC, Roediger HL. Test-enhanced learning in medical education. Med Educ. 2008;42(10):959-66.
13. Brame CJ, Biel R. Test-enhanced learning: the potential for testing to promote greater learning in undergraduate science courses. CBE Life Sci Educ. 2015;14(2):14:es4.
14. Castillo J, Gomar C, Higueras E, Gallart A. Checklist-based scores overestimate competence in CPR compared with recording strips of manikins in BLS courses. Resuscitation. 2017;114:e17.
15. David Gettman presented The Application of Generalized Estimating Equations (GEE) in Pharmacy Practice Research: Methods, Cases, and Statistical Analysis at: Special Presentation during required course entitled Pharmacy Practice Research Methods & Statistics (PHARM747); Apri, 1998; Albuquerque, New Mexico. https://doi.org/10.13140/RG.2.2.10322.06084.
16. Ballinger GA. Using Generalized Estimating Equations for Longitudinal Data Analysis. Organizational Research Methods. 2004;7(2), 127-150.
17. Kromann C, Jensen M, Ringsted C. Test-enhanced learning may be a gender-related phenomenon explained by changes in cortisol level. Medical Education. 2011;45:192–199.
18. Rowland CA. The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. Psychological Bulletin. 2014;140(6),1432–1463.
19. Lakomek F, Lukas RP, Brinkrolf P, et al. Real-time feedback improves chest compression quality in out-of-hospital cardiac arrest: A prospective cohort study. Plos One. 2020;15(4):e0232672.
20. Downing M, Sakarcan E, Quinn K. The History of Cardiopulmonary Resuscitation and Where We Are Today. Hearts. 2025;6(1):8.
21. Jansen JJ, Berden HJ, van der Vleuten CP, Grol RP, Rethans J, Verhoeff CP. Evaluation of cardiopulmonary resuscitation skills of general practitioners using different scoring methods. Resuscitation. 1997;34(1):35-41.

Figures

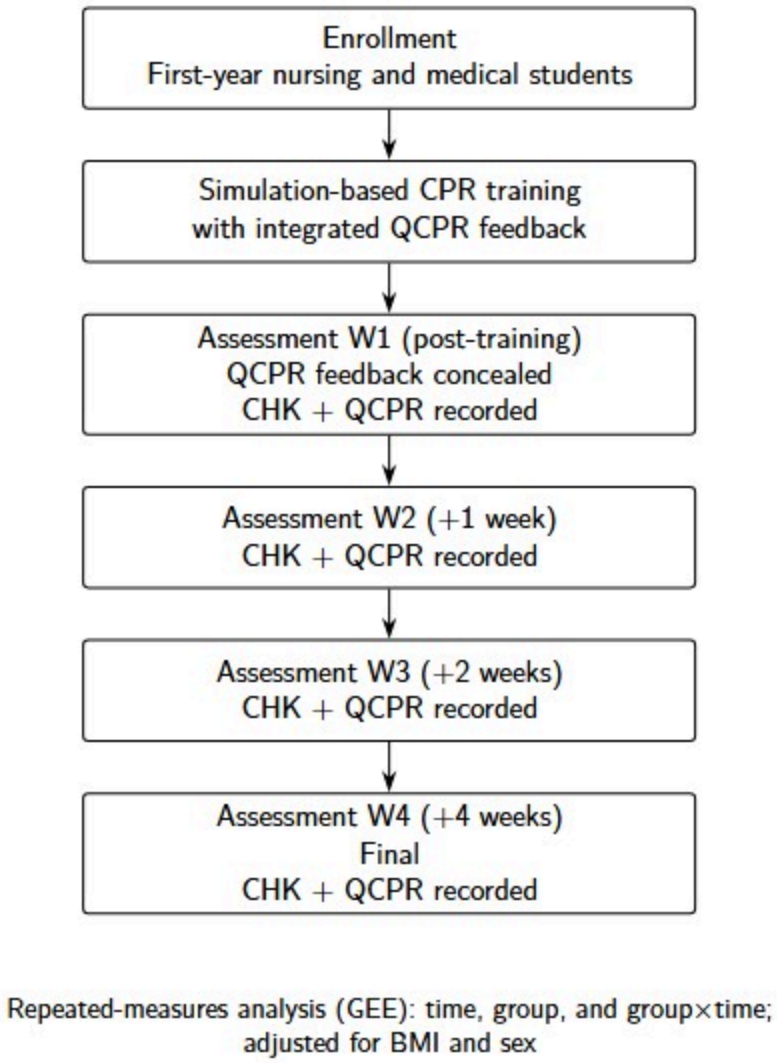


Figure 1

Study Workflow and Assessment Schedule

CPR: cardiopulmonary resuscitation; QCPR: device-derived quality cardiopulmonary resuscitation; W1: immediately post-training; W2: week 1; W3: week 2; W4: week 4; GEE: generalized estimating equations; BMI: body mass index; CHK: checklist score.

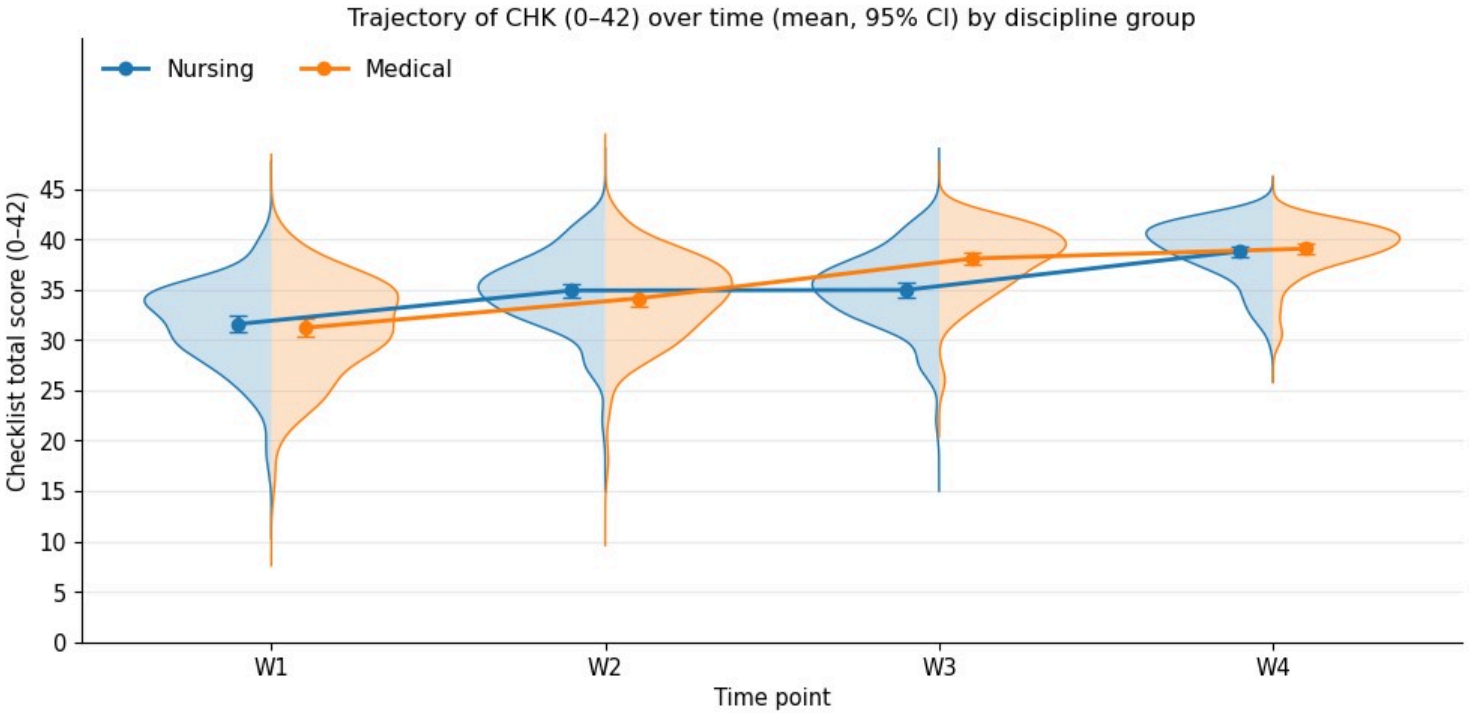


Figure 2

Trajectory of Checklist Performance across Four Assessments by Discipline Group

Points indicate mean CHK at each time point and error bars represent 95%CI. Shaded distributions depict the score density within each group at each assessment.

W1: immediately post-training; W2: week 1; W3: week 2; W4: week 4; CHK: checklist score; 95%CI: 95% Confidence Interval.

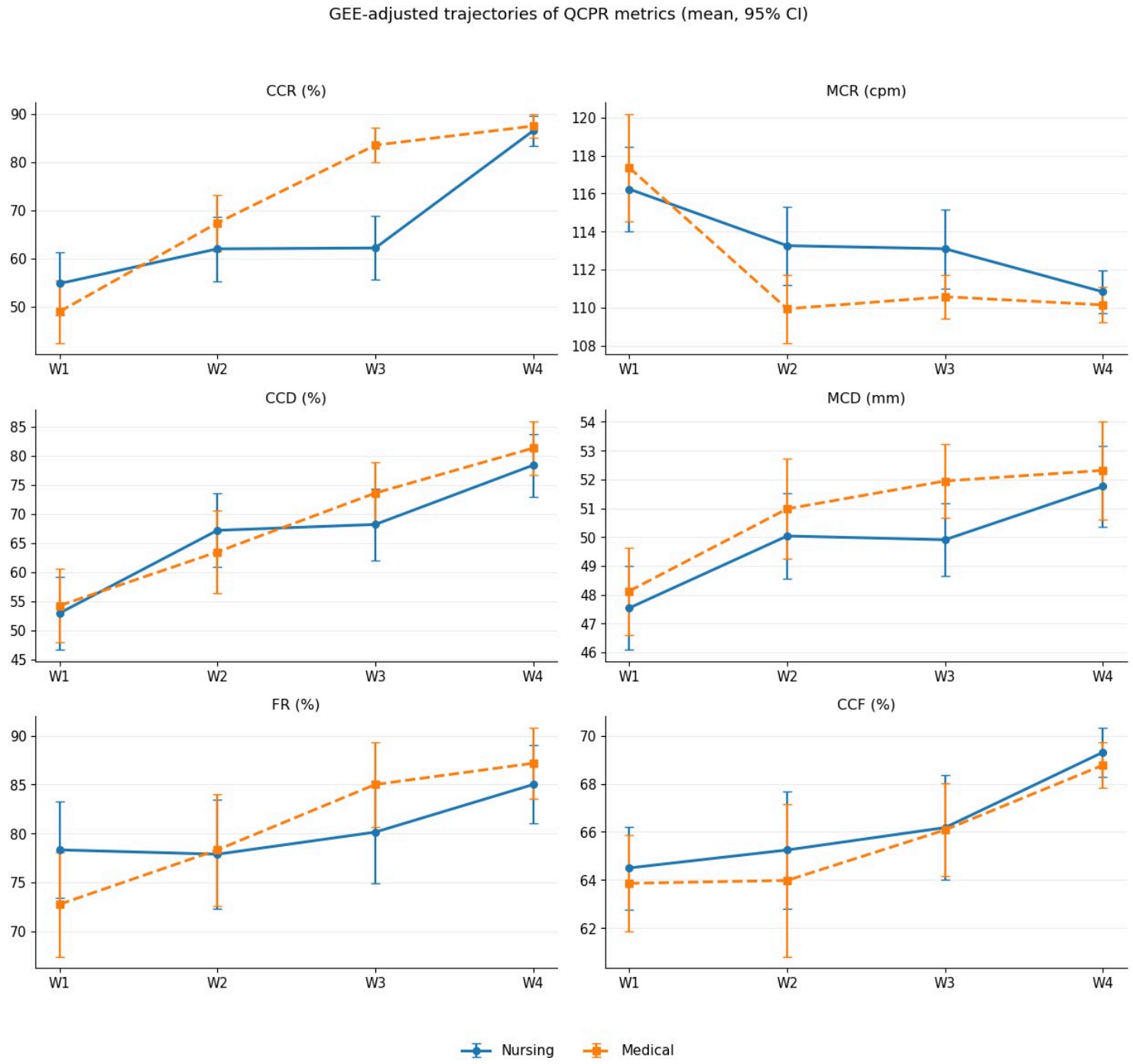


Figure 3

GEE-adjusted Trajectories of Device-derived Quality Cardiopulmonary Resuscitation Metrics over Four Assessments by Discipline Group

Panels show compression rate-in-range (CCR, % within 100–120/min), mean compression rate (MCR, cpm), compression depth-in-range (CCD, % within 50–60 mm), mean compression depth (MCD, mm), full release (FR, % complete recoil), and chest compression fraction (CCF, % time in compressions). Points represent model-based adjusted means and error bars indicate 95% confidence intervals from Gaussian generalized estimating equations (GEE) with an autoregressive (1) working correlation structure,

adjusted for sex and body mass index. Solid lines denote nursing students and dashed lines denote medical students; real-time manikin feedback was concealed during assessments.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 5.SupplementalDigitalContent.docx
- 4d.figS1.jpeg
- 6.TableS3.docx