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.

Just-in-time simulation rehearsal accelerates skill acquisition in oral flexible bronchoscopic intubation by anaesthesia residents: a propensity-score–matched cohort study

Mi Li
Chongqing Cancer Hospital
XiXi Tang
Chongqing Cancer Hospital
Chengfu Zhou
Chongqing Cancer Hospital
Yaohui Huang
Chongqing Cancer Hospital
Qi Chen
mozzie0518@cqu.edu.cn
Chongqing Cancer Hospital
Xianzhang Zeng
Chongqing Cancer Hospital

Research Article

Keywords: Just-in-time training, Simulation-based education, Fiberoptic bronchoscope-guided endotracheal intubation, Learning curve, Propensity score matching

Posted Date: August 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10589932/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

Background: Fiberoptic bronchoscope-guided endotracheal intubation is a core but difficult skill for anaesthesia residents, and the optimal way to schedule simulation practice is uncertain. Just-in-time (JIT) training—a brief structured rehearsal on an airway model performed immediately before the clinical procedure—has improved performance for other bedside procedures, but has not been evaluated for bronchoscopic intubation. We compared residents trained with JIT rehearsal versus conventional block-style simulation training.

Methods: In this single-centre retrospective cohort study, residents who underwent JIT rehearsal (introduced in 2024) were compared with residents trained conventionally. From a pool of 61 residents (20 JIT, 41 conventional candidates), 1:1 propensity-score matching on four baseline covariates (seniority, prior bronchoscopic intubations, prior bronchoscopic training, prior count as procedural assistant) yielded 20 matched pairs (40 residents). Each resident was assessed over four supervised oral bronchoscopic intubations. The primary outcome was technical quality (blinded global rating scale, GRS); secondary outcomes were a procedure-specific checklist, intubation time, NASA Task Load Index (NASA-TLX), success, attempts, lowest SpO_2 and haemodynamic change. Because each resident contributed four repeated assessments, linear mixed-effects models tested the group×assessment interaction (learning rate) and the fourth-assessment (endpoint) contrast; analyses were exploratory.

Results: Groups were comparable at the first assessment (GRS 78.5 vs 78.2). JIT residents improved faster (GRS gain 2.62 vs 1.72 points per assessment; slope difference 0.90, 95% CI 0.50–1.29; group×assessment interaction $p<0.001$) and by the fourth assessment scored 2.9 GRS points higher (95% CI 1.36–4.44, $p<0.001$). The checklist showed the same pattern (interaction $p<0.001$). Time-averaged group differences were small (GRS +1.4, $p=0.04$) and attenuated after adjustment for the residual covariate imbalance ($p=0.06$). Intubation time and NASA-TLX fell substantially in both groups with no between-group difference (interaction $p=0.72$ and 0.44). Overall success was high in both groups (98% [78/80] vs 92% [74/80], $p=0.28$), and safety measures (lowest SpO_2, heart-rate and blood-pressure change) were equivalent.

Conclusions: In this exploratory matched cohort, JIT simulation rehearsal was associated with faster acquisition of bronchoscopic-intubation quality, while efficiency, cognitive load, success and safety were comparable to conventional training. JIT rehearsal is a low-cost, embeddable adjunct that may accelerate the early learning curve; the small, single-centre effect warrants confirmation in a multicentre randomised trial.

Trial registration: The trial was registered at the Chinese Clinical Trial Registry (ChiCTR2600128222, registry URL: https://www.chictr.org.cn/showproj.html?proj=317875) in 16 July 2026.

Introduction

Fiberoptic bronchoscope-guided endotracheal intubation is a cornerstone technique for the anticipated and unanticipated difficult airway, and competence in it is an expected outcome of anaesthesia residency[1]. It is also among the harder airway skills to acquire: the technique couples fine bimanual scope manipulation with continuous visual–spatial interpretation, and learning curves derived from cumulative-sum analyses show that many supervised cases are needed before residents reach reliable, independent performance[2–3]. Because the first attempts a trainee performs are on real patients, the cost of slow or low-quality early performance is not only educational but clinical—repeated or prolonged attempts are associated with desaturation and other adverse events[4].

Simulation-based education improves procedural performance across health professions and, when paired with deliberate, distributed practice, can shorten the path to competence[5–7]. For bronchoscopic intubation specifically, task-trainer and virtual-reality programmes improve skill in the laboratory[8–10]. Most of this evidence, however, concerns how much and how realistic the practice is, rather than when it is delivered. In conventional curricula, simulation is front-loaded as block teaching—a lecture and model practice at the start of a rotation—after which trainees accumulate clinical cases without a structured rehearsal immediately beforehand. The interval between that initial practice and a given clinical case can be long, and skills decay.

Just-in-time (JIT) training reframes the timing question. Rather than massing practice at the start of training, a brief, structured rehearsal is performed on a model immediately before the clinical procedure, at the point of care. For other bedside procedures—infant lumbar puncture, neonatal and paediatric intubation, and overnight intensive-care procedures—JIT (often combined with just-in-place delivery) has improved first-attempt success and performance[11–14], consistent with deliberate-practice and mastery-learning principles in which targeted, timely repetition consolidates a fragile skill[15]. To our knowledge, JIT rehearsal has not been evaluated for bronchoscopic intubation in anaesthesia residents, where the skill is complex, the early curve is steep, and the timing of practice relative to the clinical case may matter most.

We therefore compared residents trained with JIT rehearsal versus conventional block-style simulation training on the quality, efficiency, cognitive load, success and safety of supervised oral bronchoscopic intubation. Because the JIT programme was introduced as a service change rather than randomised, we used propensity-score matching to improve baseline comparability, and we modelled the four serial assessments per resident to test not only average performance but the rate of skill acquisition. We pre-specified technical quality and its learning trajectory as the primary focus.

Materials and Methods

Study design and setting

This single-center retrospective cohort study was approved by the institutional ethics committee(CZLL2026-077-002), with a waiver of informed consent granted due to its retrospective

design. This study was conducted in accordance with the Declaration of Helsinki and was registered at the Chinese Clinical Trial Registry (ChiCTR2600128222, registry URL: https://www.chictr.org.cn/showproj.html?proj=317875) in 16 July 2026. Principal investigator was Qi Chen. It was a study of anaesthesia residents performing supervised fiberoptic bronchoscope-guided endotracheal intubation at Chongqing University Affiliated Cancer Hospital. A JIT rehearsal programme was introduced in 2024; residents trained under this programme were compared with residents who had completed conventional training.

Participants and eligibility

Eligible participants were anaesthesia resident ssince 2024 who performed supervised oral bronchoscopic intubation during the study period and had complete assessment documentation.

Exclusion criteria: (1) Physicians who had received standardized systematic fiberoptic bronchoscope training and independently completed more than 10 successful fiberoptic bronchoscope-guided orotracheal intubations; (2) Participants with incomplete video records.

Training interventions

JIT group

Standardized simulation training for fiberoptic bronchoscope-guided orotracheal intubation was conducted on mannequins within 1 hour before formal clinical intubation under the guidance of chief instructors. Each trainee was required to complete at least one successful intubation or receive 10 minutes of standardized training.

traditional (control) group

Residents received block-style training—centralised instruction and model practice at the start of the rotation—followed by supervised clinical practice without an immediate pre-procedure rehearsal.

Patients with potential difficult airways were selected for clinical training, meeting at least one of the following criteria: mouth opening less than two finger breadths, neck mobility less than 90°, short and thick neck, obstructive sleep apnea syndrome, history of snoring, micrognathia, thyromental distance less than 6.5 cm, facial deformity, and obesity (body mass index > 26 kg/m²).

Intubation was terminated and defined as a failure if any of the following situations occurred: total intubation duration exceeding 8 minutes, more than 3 intubation attempts, peripheral oxygen saturation dropping below 85%, or blurred visual field caused by bleeding[16]. Experienced attending anesthesiologists would take over the procedure immediately once intubation failed.

All enrolled patients were preoperatively evaluated and selected by an expert team consisting of attending anesthesiologists, chief residents and chief physicians. All patients underwent fiberoptic bronchoscope-guided orotracheal intubation under general anesthesia. The entire procedure was

recorded by a dual-camera wireless fiberoptic bronchoscope system: one camera captured the endoscopic view, and the other recorded the operator's maneuvers and real-time vital sign monitoring data.

After admission, electrocardiography, pulse oximetry and non-invasive blood pressure monitoring were established for all patients. Atropine 0.3–0.5 mg was intravenously administered to reduce airway secretions if no contraindications existed. Anesthesia induction was performed with midazolam 1–2 mg, propofol 2–3 mg/kg, sufentanil 0.3–0.5 μg/kg and rocuronium 0.6–0.9 mg/kg. Fiberoptic bronchoscope-guided orotracheal intubation was initiated after full onset of anesthetic drugs.

Assessments and outcomes

Each resident was assessed on four supervised oral bronchoscopic intubations. Performance was scored by Qi Chen using two instruments: a global rating scale (GRS) of overall technical quality and a procedure-specific checklist, both scored so that higher values indicate better performance[17–18]. The primary outcome was the GRS and its learning trajectory across the four assessments. Secondary outcomes were: the procedure-specific checklist; intubation time (seconds; lower is better); perceived workload by the NASA Task Load Index (NASA-TLX), a validated measure of cognitive load in airway and procedural tasks (lower is better)[19–20]; intubation success; number of attempts; and safety measures—lowest SpO_2 during the procedure and the change in heart rate (ΔHR) and blood pressure (ΔBP) from before to after intubation.

Propensity-score matching

To improve baseline comparability between the non-randomised groups, we estimated each resident's propensity for JIT training by logistic regression on four pre-specified covariates: seniority (years), prior number of bronchoscopic intubations, prior bronchoscopic training (yes/no), and prior count as a procedural assistant[21]. From the complete pool of 61 residents (20 JIT and 41 conventional candidates), JIT residents were matched 1:1 to conventional residents on the propensity score, yielding 20 matched pairs (40 residents). Balance was assessed using absolute standardised mean differences (SMD) before and after matching, with SMD < 0.1 taken to indicate negligible imbalance, and propensity-score overlap was examined in the full pre-match pool[22].

Statistical analysis

Per-resident summaries (mean of the four assessments) were compared between groups using the Wilcoxon rank-sum test for continuous outcomes; overall success (all 4×20 intubations per group) was compared using Fisher's exact test.

Because each resident contributed four repeated assessments, we additionally fitted linear mixed-effects models with a random intercept for resident. For each continuous outcome we modelled training group, assessment, and their interaction. Assessment was modelled as a continuous term to estimate

the learning rate (per-assessment slope) and the group×assessment interaction (whether the rate of improvement differed between groups), and as a factor to estimate timepoint-specific contrasts; the fourth-assessment (endpoint) between-group difference and the time-averaged group difference were obtained as estimated marginal means with Kenward–Roger degrees of freedom. A sensitivity analysis additionally adjusted the group effect for prior assist count, the covariate with residual imbalance after matching.

Two-sided $p < 0.05$ was considered significant. Given the modest sample and multiple outcomes, all analyses are exploratory and were not adjusted for multiplicity; estimates are reported with 95% confidence intervals and interpreted by effect size and consistency of direction rather than by significance thresholds alone. Analyses used R 4.3.1.

Results

Participants and matching

The complete pre-match pool comprised 61 residents (20 JIT, 41 conventional candidates). A CONSORT flow diagram describing the progression of subjects through the study is presented in Fig. 1. Propensity scores overlapped well across groups (range ≈ 0.25–0.75), supporting comparability. After 1:1 matching, 20 JIT residents were paired with 20 traditional residents (40 residents; 160 supervised intubations, 80 per group). Matching improved balance on the three imbalanced covariates—prior assist count (SMD 0.52→0.26), prior bronchoscopic training (0.38→0.10), and prior bronchoscopic intubations (0.17→0.14)—while seniority, already balanced, changed minimally (0.03→0.13) (Table 1, Fig. 6). Residual imbalance was concentrated in prior assist count (SMD 0.26) and was addressed in a sensitivity analysis.

Table 1
Covariate balance before and after propensity-score matching (matched-set median [Q1, Q3] or n (%); absolute SMD).

Covariate	Traditional (matched)	Just-in-time	SMD before	SMD after
Seniority (yr)	2.0 (2.0, 3.0)	2.0 (1.8, 3.0)	0.034	−0.133
Prior bronchoscopic intubations	1.0 (0.0, 1.0)	0.0 (0.0, 1.0)	−0.174	−0.143
Prior bronchoscopic training (yes)	7 (35%)	6 (30%)	−0.384	−0.104
Prior assist count	2.0 (1.8, 3.0)	3.0 (2.0, 3.0)	0.519	0.260

Primary outcome: technical quality and its learning trajectory

At the first assessment the groups were essentially identical (model-based GRS 78.5 [JIT] vs 78.2 [traditional]). Both groups improved across the four assessments, but JIT residents improved faster: the per-assessment GRS gain was 2.62 points in the JIT group versus 1.72 in the traditional group, a slope difference of 0.90 points per assessment (95% CI 0.50–1.29), and the group×assessment interaction was significant ($p < 0.001$) (Fig. 2). The gap widened over time, so that by the fourth assessment JIT residents scored 2.9 GRS points higher (95% CI 1.36–4.44, $p < 0.001$). Averaged across all four assessments, the GRS advantage was small (+ 1.4 points, 95% CI 0.06–2.76, $p = 0.041$; descriptive medians 82.4 vs 81.6, $p = 0.044$; Table 2) and attenuated to a non-significant trend after adjustment for the residual prior-assist-count imbalance (+ 1.2 points, 95% CI − 0.07–2.55, $p = 0.063$). The robust signal was therefore the trajectory—a steeper within-resident learning curve—rather than a uniform shift in level.

Table 2
Intubation outcomes: just-in-time vs traditional training (per-trainee mean over four assessments; median [Q1, Q3]; Wilcoxon rank-sum, except success by Fisher exact).

Outcome	Traditional	Just-in-time	p
GRS score	81.6 (79.4, 82.8)	82.4 (80.9, 84.5)	0.044
Checklist score	31.0 (28.5, 31.6)	31.8 (30.4, 33.6)	0.088
Intubation time (s)	158.1 (139.8, 224.8)	156.8 (132.0, 185.2)	0.598
NASA-TLX (cognitive load)	51.8 (47.8, 57.4)	50.8 (38.7, 56.4)	0.512
Lowest SpO_2 (%)	97.5 (96.2, 98.6)	98.0 (96.8, 99.0)	0.551
Attempts	1.2 (1.0, 1.3)	1.0 (1.0, 1.2)	0.268
ΔHR (post − pre)	14.8 (12.2, 20.7)	16.0 (12.7, 18.0)	0.978
ΔBP (post − pre)	1.4 (− 2.3, 5.8)	4.9 (0.1, 7.8)	0.113
Success rate (all intubations)	92% (74/80)	98% (78/80)	0.276

Secondary outcomes: quality checklist, efficiency, cognitive load

The procedure-specific checklist mirrored the GRS: groups started equal and the JIT group improved faster (slope difference 0.80 per assessment, 95% CI 0.45–1.14; interaction $p < 0.001$), reaching a 2.45-point advantage by the fourth assessment (95% CI 1.03–3.87, $p = 0.001$); the time-averaged difference was small (descriptive 31.8 vs 31.0, $p = 0.088$; Table 2).

In contrast, efficiency and workload improved markedly in both groups with no between-group difference. Intubation time fell by roughly 35–36 seconds per assessment in each group (group×assessment interaction $p = 0.72$; time-averaged 156.8 vs 158.1 s, $p = 0.60$). NASA-TLX declined

comparably in both groups (interaction p = 0.44; time-averaged 50.8 vs 51.8, p = 0.51) (Fig. 2, Fig. 3, Table 2). Both arms thus learned to intubate faster and with lower perceived workload at the same rate; the JIT advantage was specific to rated technical quality.

Success and safety

Overall intubation success was high in both groups, with no significant difference (JIT 98% [78/80] vs traditional 92% [74/80]; Fisher p = 0.28), and success rose toward a ceiling across timepoints in both arms (Fig. 4). The number of attempts did not differ (1.0 vs 1.2, p = 0.27). Safety measures were equivalent: lowest SpO_2 (98.0% vs 97.5%, p = 0.55), ΔHR (16.0 vs 14.8, p = 0.98) and ΔBP (4.9 vs 1.4, p = 0.11) (Fig. 5, Table 2). JIT rehearsal conferred no additional safety benefit but was not associated with any decrement.

Discussion

In this matched cohort of anaesthesia residents learning fiberoptic bronchoscope-guided endotracheal intubation, just-in-time simulation rehearsal was associated with a faster rate of improvement in rated technical quality: the two groups began at the same level and JIT residents pulled ahead over four supervised cases on both the global rating scale and the procedure checklist (group×assessment interaction p < 0.001), reaching a clinically modest but statistically clear advantage by the fourth assessment. Efficiency (intubation time), cognitive load (NASA-TLX), success and patient-safety measures improved with practice but did so equally in both groups. The most defensible reading is therefore narrow and specific: JIT rehearsal appears to accelerate the early quality learning curve without changing the procedure's efficiency, success or safety endpoints.

This pattern is what timing-based theory predicts. The benefit of JIT is hypothesised to come from rehearsing a fragile skill immediately before performance, minimising decay between practice and execution[11,15]. It is coherent that such a mechanism would most affect the components of performance that are most sensitive to recent rehearsal—technique and procedural completeness, captured by the GRS and checklist—while end-state metrics such as success and saturation, which are high and near-ceiling for supervised cases in both arms, are less able to discriminate. Our findings extend JIT evidence from lumbar puncture and neonatal/paediatric intubation[11–13] to a more complex bimanual airway skill, and they are consistent with the broader simulation literature in which the scheduling and deliberate structuring of practice—not merely its quantity—shape acquisition[5–7].

The clinical and educational implications are practical. JIT rehearsal is inexpensive, brief and embeddable into existing clinical workflow, and our data suggest it may help residents reach high-quality technique sooner. Because early bronchoscopic attempts occur on real patients, accelerating the quality curve—even without a measurable change in success during supervised practice—has plausible value for patient experience and for trainee confidence, and may shorten the supervised period required before

independent practice. At the same time, the absence of any safety or success advantage should be communicated honestly: JIT is a means of more efficient learning, not a safety intervention in itself.

Strengths

Several features strengthen inference from a non-randomised design. Propensity-score matching on four pre-specified covariates improved baseline comparability and made the residual imbalance explicit rather than hidden. The four serial assessments per resident allowed a within-person analysis of the learning trajectory using mixed-effects models, which is both more powerful and more informative than comparing single endpoints—and it is precisely the trajectory, not a level shift, that proved robust. The convergence of two independent quality instruments (GRS and checklist) on the same pattern further supports the finding.

Limitations

This study has important limitations and its results should be regarded as exploratory and hypothesis-generating. First, the design is single-centre and retrospective, with non-random allocation; unmeasured confounding (for example, secular improvements in teaching that coincided with the 2024 introduction of JIT) cannot be excluded. Second, the sample is small (20 matched pairs), so estimates are imprecise and the time-averaged quality advantage was fragile—significant in the primary analysis but attenuating to a trend after adjustment for residual imbalance; we therefore emphasise the within-person interaction, which is less vulnerable to baseline differences. Third, residual imbalance remained on prior assist count (SMD 0.26) despite matching; although we adjusted for it, finer control was limited by sample size. Fourth, success and lowest SpO_2 showed ceiling effects under supervision and cannot exclude differences that might emerge in unsupervised or genuinely difficult airways. Finally, the absolute quality difference was small (≈ 1 GRS point on average; ≈3 points at the endpoint), and its clinical importance is uncertain.

Conclusions

Among anaesthesia residents learning fiberoptic bronchoscope-guided endotracheal intubation, just-in-time simulation rehearsal was associated with a faster gain in rated technical quality over four supervised cases, while intubation time, cognitive load, success and safety were comparable to conventional block-style training. JIT rehearsal is a low-cost, workflow-compatible adjunct that may accelerate the early learning curve. Given the modest, single-centre, exploratory nature of these findings, a multicentre randomised trial—powered for the learning-rate interaction and including unsupervised or difficult-airway performance and skill retention—is warranted before broad adoption.

Abbreviations

JIT, just-in-time; GRS, global rating scale; NASA-TLX, NASA Task Load Index; OSATS, Objective Structured Assessment of Technical Skills; SMD, standardised mean difference; CI, confidence interval; ΔHR,

change in heart rate; ΔBP, change in blood pressure; SpO_2, peripheral oxygen saturation.

Declarations

Ethics approval and consent to participate: Ethical approval was obtained from Chongqing University Affiliated Cancer Hospital(CZLL2026-077-002) and Given its retrospective design, the study obtained a waiver of informed consent from our hospital's ethics review board.

Consent for publication: All authors read and approved the final manuscript.

Availability of data and materials: The datasets analysed are available from the corresponding author on reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding : This work was supported by National High Level Hospital Clinical Research Funding and National Cancer Center Climbing Fund (No. NCC202422001) and the Key Joint Medical Research Project of Chongqing Municipal Science and Technology Commission (Grant No. 2025ZDXM014). Chongqing Municipal Health Commission Traditional Chinese Medicine Research Project(2026WSJK143) and Scientific and Technological Research Project of Chongqing Municipal Education Commission (Grant No. KJQN202400108).

Authors' contributions: Study concept and design: Q.C, X.Z. Acquisition of data: X.T, M.L, C.Z, Y.H. Manuscript drafting: X.T, M.L, Q.C. Methodology and software: X.T, M.L, Study supervision: X.Z, Q.C. All authors read and approved the final manuscript.

Acknowledgments: The authors have no acknowledgements.

References

1. Baker P. Preparedness and education in airway management. Anesthesiol Clin. 2015;33(2):381–95.
2. Smith JE, Jackson AP, Hurdley J, Clifton PJ. Learning curves for fibreoptic nasotracheal intubation when using the endoscopic video camera. Anaesthesia. 1997;52(2):101–6.
3. Weil G, Motamed C, Biau DJ, Guye ML. Learning curves for three specific procedures by anesthesiology residents using the learning curve cumulative sum (LC-CUSUM) test. Korean J Anesthesiol. 2017;70(2):196–202.
4. Takahashi J, Goto T, Fujitani S, et al. Association of airway obstruction with first-pass success and intubation-related adverse events in the emergency department: multicenter prospective observational studies. Front Med (Lausanne). 2023;10:1199750.
5. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA. 2011;306(9):978–88.

6. McGaghie WC, Barsuk JH, Wayne DB, Issenberg SB. Powerful medical education improves health care quality and return on investment. Med Teach. 2024;46(1):46–58.
7. Cook DA, Hamstra SJ, Brydges R, et al. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. Med Teach. 2013;35(1):e867–98.
8. Melvin MT, Siddiqui NT, Wild E, Parotto M, Perelman VS, You-Ten KE. Achieving Competency in Fiber-Optic Intubation Among Resident Physicians After Higher- Versus Lower-Fidelity Task Training: A Randomized Controlled Study. Anesth Analg. 2023;137(1):200–8.
9. Cailleau L, Geeraerts T, Minville V, et al. Is there a benefit for anesthesiologists of adding difficult airway scenarios for learning fiberoptic intubation skills using virtual reality training? A randomized controlled study. PLoS ONE. 2023;18(1):e0281016.
10. Duffy CC, Bass GA, Yi W, Rouhi A, Kaplan LJ, O'Sullivan E. Teaching Airway Management Using Virtual Reality: A Scoping Review. Anesth Analg. 2024;138(4):782–93.
11. Kessler D, Pusic M, Chang TP, et al. Impact of Just-in-Time and Just-in-Place Simulation on Intern Success With Infant Lumbar Puncture. Pediatrics. 2015;135(5):e1237–46.
12. Gizicki E, Assaad MA, Massé É, Bélanger S, Olivier F, Moussa A. Just-In-Time Neonatal Endotracheal Intubation Simulation Training: A Randomized Controlled Trial. J Pediatr. 2023;261:113576.
13. McKay J, Wasserman M, Monuteaux MC, Hirsch AW, Nagler J. Just-in-time procedural training for pediatric emergency medicine trainees: A randomized educational interventional trial. AEM Educ Train. 2023;7(3):e10886.
14. Rabinowitz R, Drake CB, Talan JW, et al. Just-in-Time Simulation Training to Augment Overnight ICU Resident Education. J Grad Med Educ. 2024;16(6):713–22.
15. Petrosoniak A, Lu M, Gray S, et al. Perfecting practice: a protocol for assessing simulation-based mastery learning and deliberate practice versus self-guided practice for bougie-assisted cricothyroidotomy performance. BMC Med Educ. 2019;19(1):100.
16. Gueret G, Billard V, Bourgain JL. Fibre-optic intubation teaching in sedated patients with anticipated difficult intubation. Eur J Anaesthesiol. 2007;24(3):239–44.
17. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg. 1997;84(2):273–8.
18. Hopmans CJ, den Hoed PT, van der Laan L, et al. Assessment of surgery residents' operative skills in the operating theater using a modified Objective Structured Assessment of Technical Skills (OSATS): a prospective multicenter study. Surgery. 2014;156(5):1078–88.
19. Ambardekar AP, Rosero EB, Bhoja R, et al. A Randomized Controlled Trial Comparing Learners' Decision-making, Anxiety, and Task Load During a Simulated Airway Crisis Using Two Difficult Airway Aids. Simul Healthc. 2019;14(2):96–103.
20. Leppink J. Cognitive load theory: Practical implications and an important challenge. J Taibah Univ Med Sci. 2017;12(5):385–91.

21. Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivar Behav Res. 2011;46(3):399–424.
22. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med. 2009;28(25):3083–107.

Figures

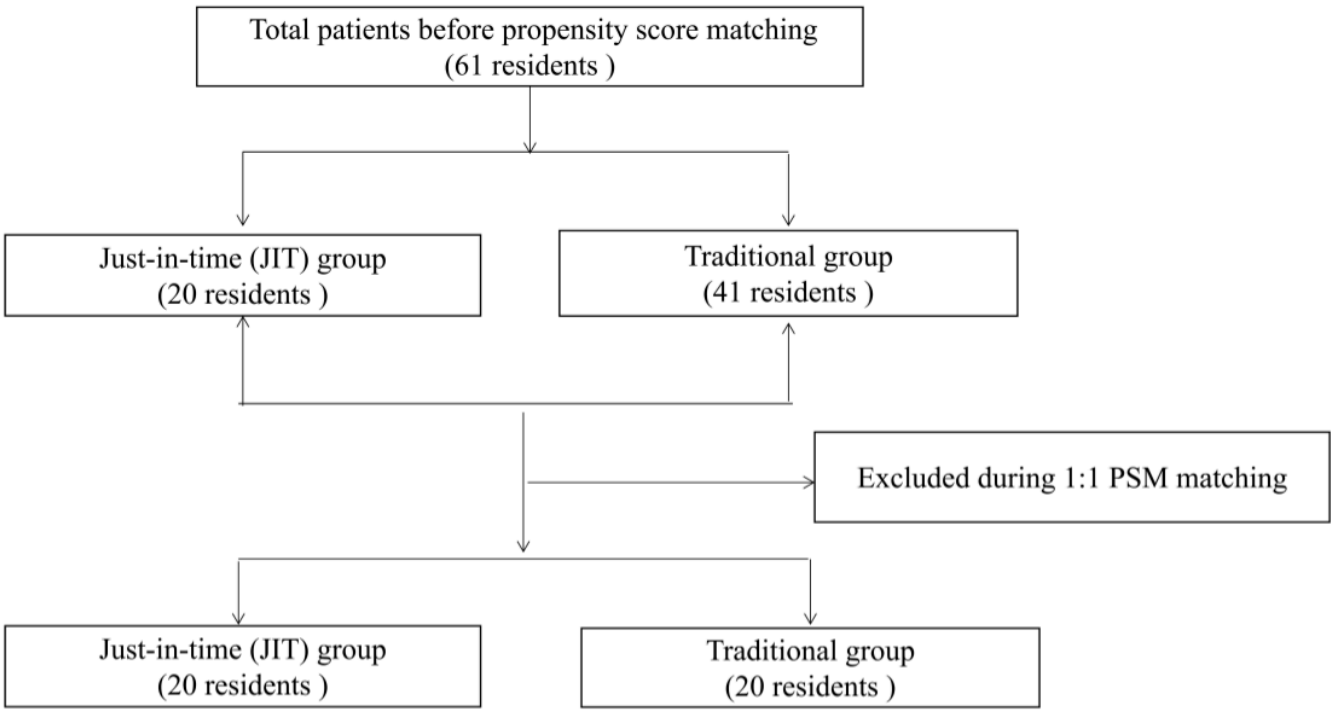


Figure 1

CONSORT flow diagram describing the progression of subjects through the study.

JIT: Just-in-time, PSM: Participants and matching

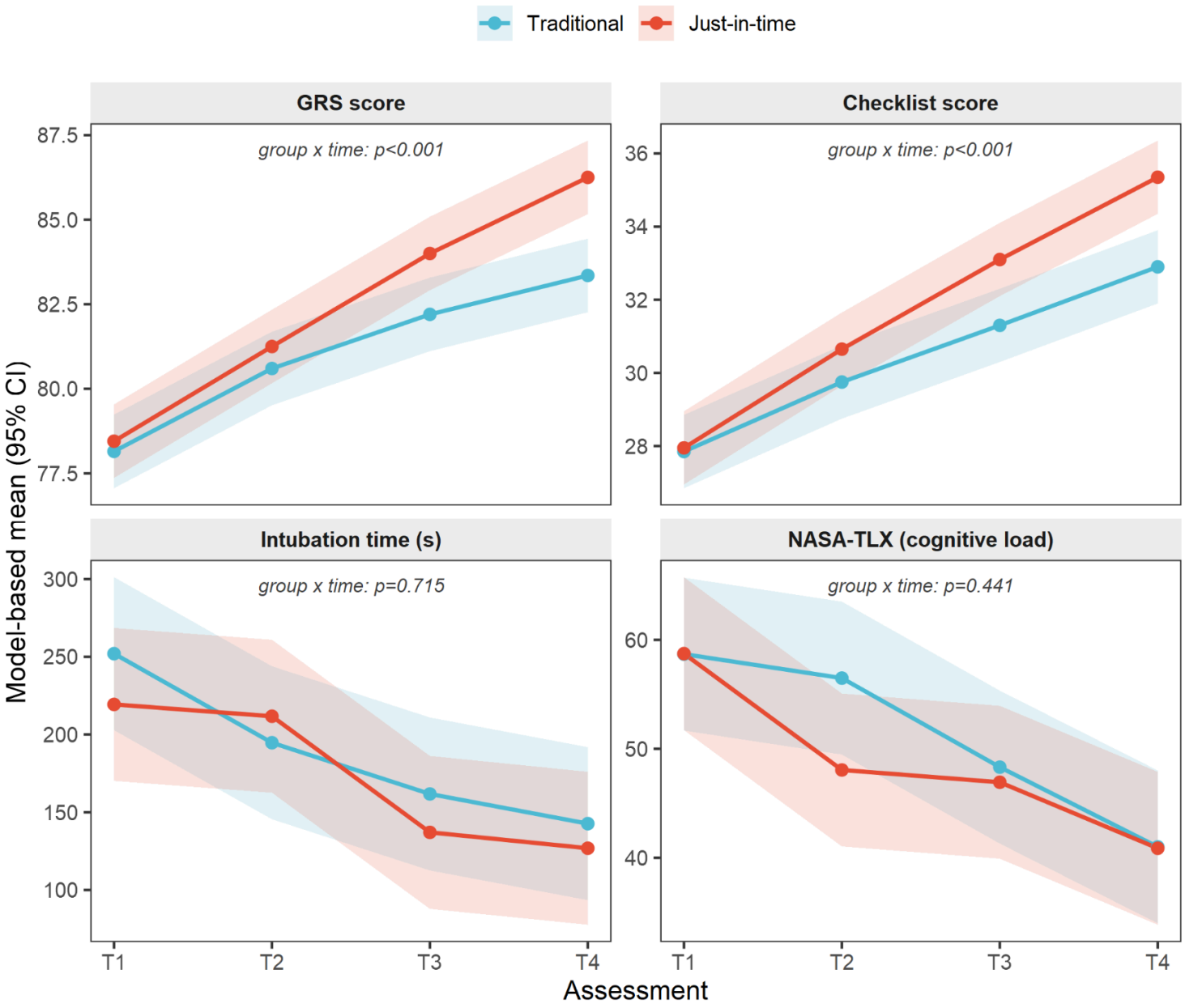


Figure 2

Model-based learning curves over four supervised assessments.

Estimated marginal means (95% CI) from linear mixed-effects models for the global rating scale (GRS), procedure checklist, intubation time and NASA-TLX, by training group (red, just-in-time; blue, traditional). The group×assessment interaction p-value is annotated in each panel. Quality measures (GRS, checklist) diverge in favour of the JIT group (interaction p<0.001), whereas intubation time and cognitive load fall in parallel (interaction p=0.72 and 0.44).

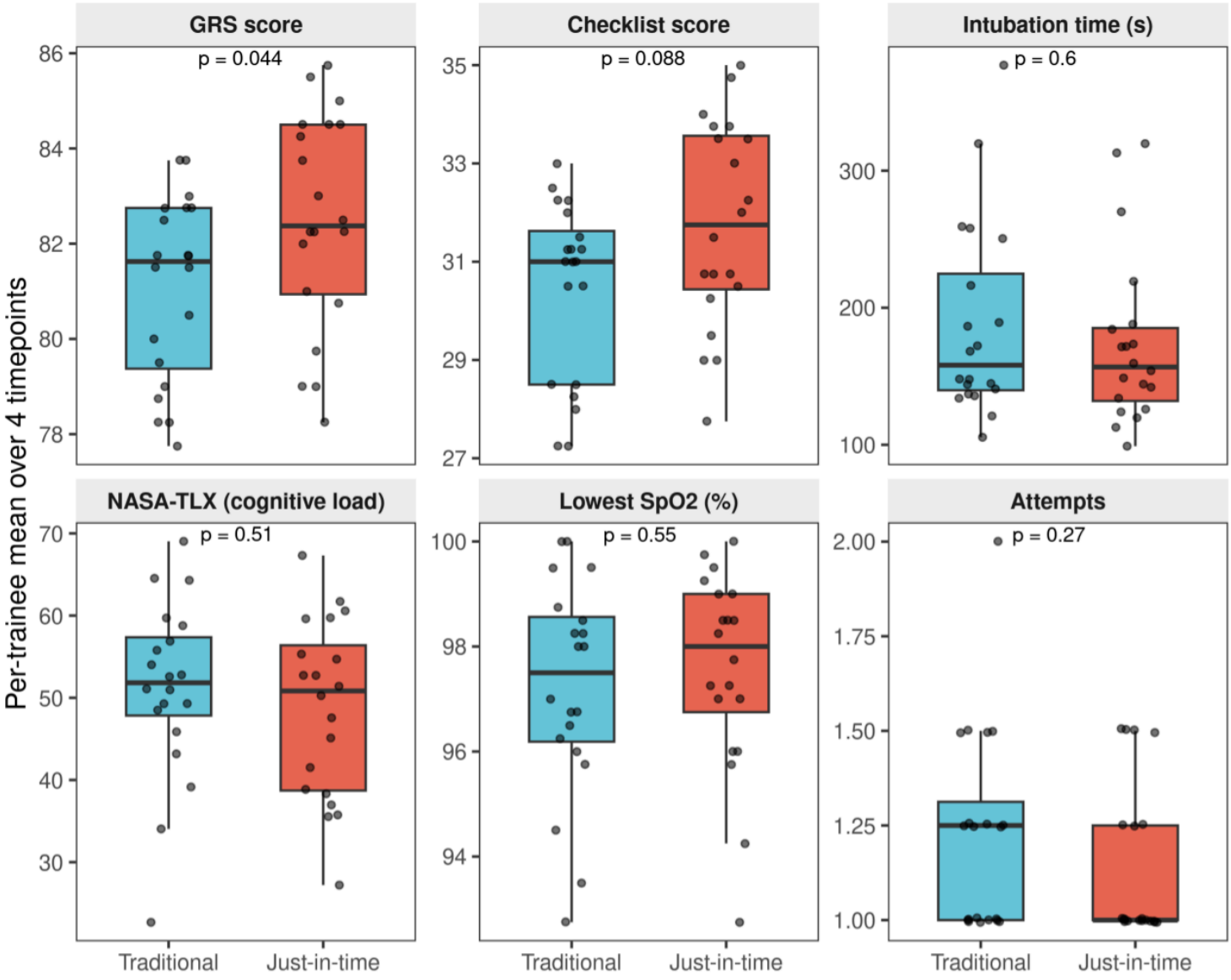


Figure 3

Per-trainee aggregate outcomes by group. Box plots of each resident's mean over the four assessments for six outcomes, with Wilcoxon p-values.

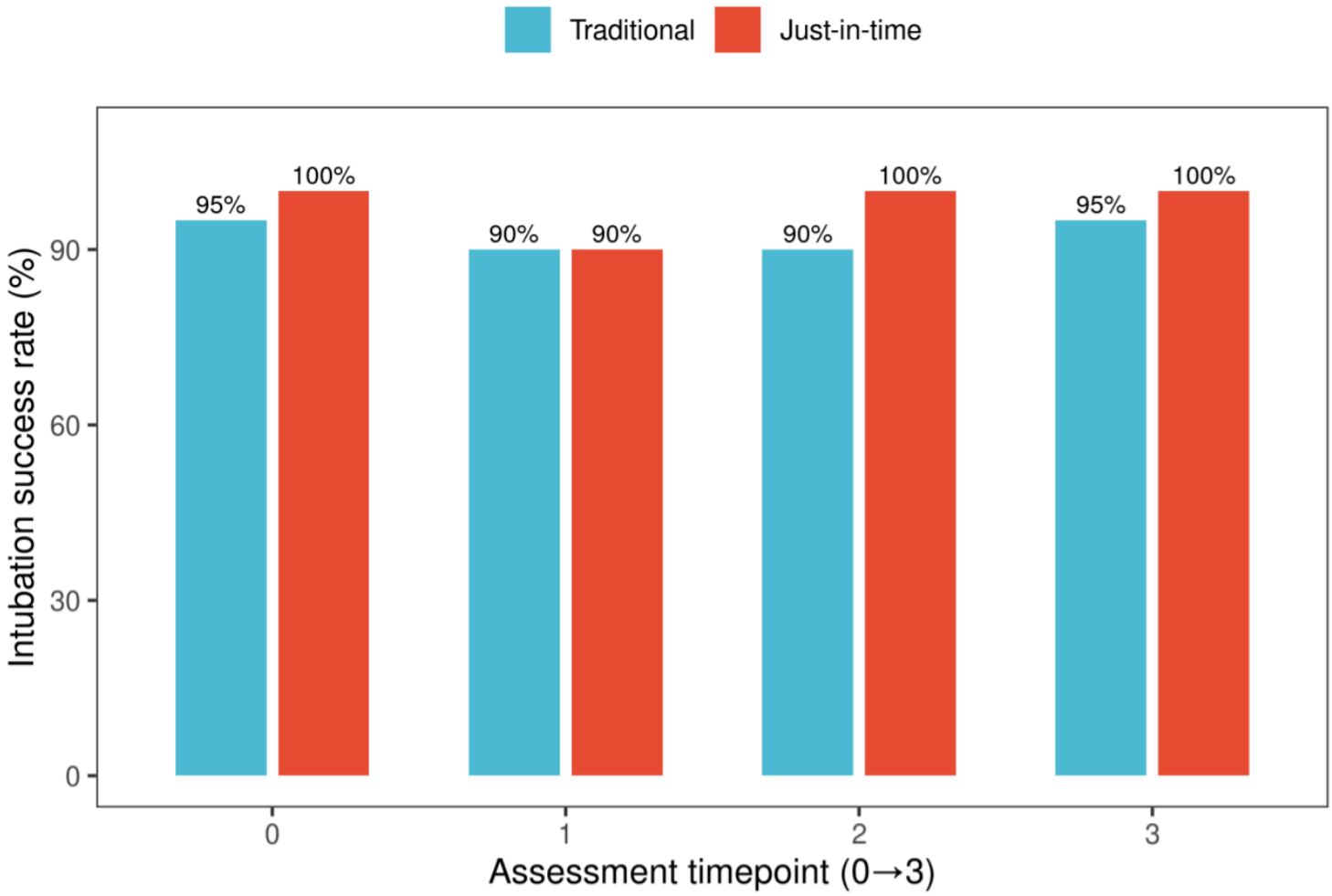


Figure 4

Intubation success rate by assessment and group.

Success was high and near-ceiling in both groups across timepoints (overall 98% vs 92%, Fisher p=0.28).

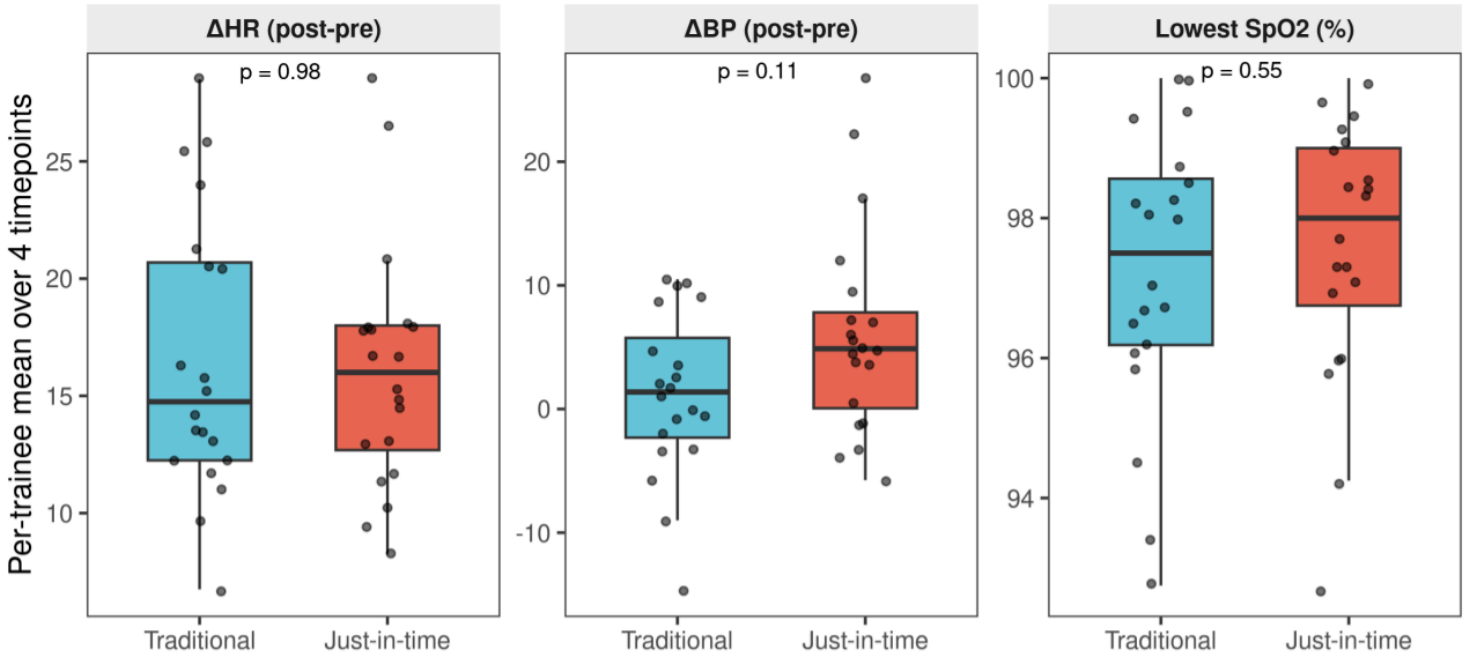


Figure 5

Safety and haemodynamic stability.

Per-trainee lowest SpO_2, ΔHR and ΔBP by group; no between-group difference.

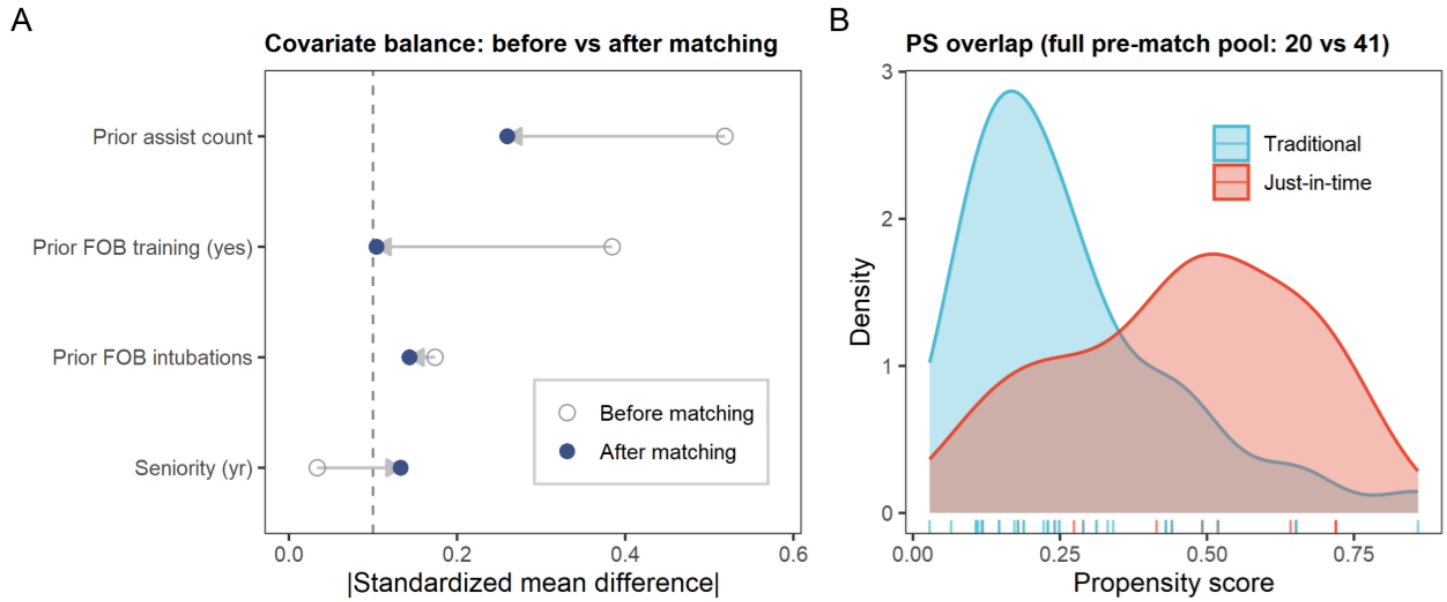


Figure 6

Propensity-score matching quality.

(A) Love plot of absolute standardised mean differences for the four covariates before (open) and after (filled) matching, relative to the 0.1 threshold (dashed). (B) Propensity-score overlap in the complete pre-

match pool (20 JIT vs 41 conventional candidates).