

Evaluating the impact of in situ simulation on interprofessional healthcare team training: a randomised controlled trial

Maria do Céu Campos

mariaceucampos@gmail.com

Faculty of Medicine, University of Porto

Jorge Jesus Silva

Department of Anaesthesiology, Unidade Local de Saúde de Matosinhos, Hospital Pedro Hispano

Rita Resende

Department of Anaesthesiology, Unidade Local de Saúde de Matosinhos, Hospital Pedro Hispano

Abel Nicolau

RISE-Health, Faculty of Medicine, University of Porto

Research Article

Keywords: in situ simulation, interprofessional education, simulation-based team training, perioperative care, local anaesthetic systemic toxicity, debriefing, randomised controlled trial

Posted Date: October 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11166437/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

Simulation-based interprofessional team training is essential to patient safety, but whether the physical setting, in situ or in a dedicated centre, influences its educational effectiveness remains unclear. Published comparisons vary the scenario, equipment or facilitators alongside the setting, so the effect of location cannot be isolated, and existing randomised trials are confined to obstetrics. This trial assessed whether the physical setting alters the educational effect of a standardised interprofessional simulation, measured by clinical knowledge, collaborative competencies and perceived effectiveness.

Methods

Multicentre, parallel-group, unblinded randomised controlled trial. Ninety-eight professionals were individually randomised 1:1 to in situ training or to a simulation centre configured to replicate the clinical service, balanced by professional category and institution. Within each setting they formed seven teams of seven (two anaesthetists, four nurses, one healthcare assistant); the intervention was delivered to the team, clustering participants after randomisation. All completed a standardised session on local anaesthetic systemic toxicity, with an equivalence matrix holding eleven intervention parameters constant. The Simulation Design Scale (SDS) and the Debriefing Assessment for Simulation in Healthcare (DASH) verified equivalent delivery. Knowledge (multiple-choice questionnaire), collaborative competencies (ICCAS) and perceived effectiveness (SET-M) were exploratory outcomes; none was primary.

Results

. All 98 participants completed the study; baseline characteristics were comparable. Delivery did not differ on the SDS ($p = 0.613$; fidelity $p = 0.305$) or the DASH total ($p = 0.677$). Knowledge rose in both settings ($p < 0.001$), with no between-setting difference (in situ vs. centre 87.35% vs. 90.48%; $p = 0.103$). ICCAS scores improved in both ($p < 0.001$), with no post-session difference (80.45 vs. 78.10; $p = 0.335$). Perceived effectiveness was comparable (51.18 vs. 51.71; $p = 0.366$). Only DASH element 4 (depth of reflection) favoured the centre (6.37 vs. 6.67; $p = 0.036$). No adverse events occurred.

Conclusions

The physical setting did not alter the educational gains or the collaborative competence acquired. Not powered as an equivalence study, the trial provides no evidence of educational disadvantage in workplace-based training. This supports building regular team simulation into clinical governance, closer to practice, where latent safety threats can also be identified.

Trial registration

: ClinicalTrials.gov, NCT07346937, <https://clinicaltrials.gov/study/NCT07346937>.

Background

The complexity of perioperative care and the interprofessional interdependence across perioperative teams make teamwork a determinant of patient safety [1, 2]. Simulation-based team training has become the educational response to this requirement, and its overall effectiveness is no longer a matter of controversy. The 2024 global consensus statement and the Østergaard Declaration of the European Society of Anaesthesiology and Intensive Care (ESAIC) consider its pedagogical efficacy established and shift the discussion to the conditions of implementation [3, 4], in a European context of marked variability in access to this training and in the time allocated to it [5].

With effectiveness settled, an operational question with direct consequences remains open. It falls to institutions to decide where to train, and that decision determines the cost, the frequency that is possible and the proportion of professionals able to take part. In situ simulation, carried out in the clinical setting itself, has been presented as an effective option, but the arguments supporting that preference are of different kinds and need to be distinguished. The literature shows that in situ simulation identifies latent safety threats and process failures that a centre does not reproduce [6, 7], that it is associated with improvements in safety culture [8] and, in a systematic review, with favourable patient outcomes [9]. These findings establish its diagnostic and organisational value. None of these studies, however, compared the educational gain obtained in situ with that obtained in a dedicated centre, and none therefore supports educational superiority.

The obstacle lies not only in the small number of comparative studies but in the methodological confounding that runs through them. Most studies vary, along with the setting, the scenario, the equipment, the facilitators or the debriefing model. When several variables change at once, the investigation evaluates two educational programmes rather than two environments, and the difference observed cannot be attributed to the setting [10]. Two randomised trials have made the comparison directly, both in obstetrics. Crofts and colleagues randomised 140 midwives and obstetricians across four settings crossing location with team training and found no effect of location on knowledge acquisition [11]. Sørensen and colleagues randomised 100 professionals to obstetric anaesthesia training in situ or off-site and found no differences in knowledge, safety attitudes, motivation or stress [12].

On this basis, three opposing predictions coexist in the literature: (1) that training in the real environment is superior, because the authenticity of roles and spatial cues favours transfer [13]; (2) that the relationship between fidelity and transfer is minimal and that the gain depends on educational design and facilitation [14, 15]; and, (3) that in situ training may be less favourable to reflective learning, since performing in front of peers and senior colleagues in one's own service increases the interpersonal risk

associated with acknowledging error [16]. No study has tested them against each other while holding the scenario, the facilitation team and the instruments constant.

Without a direct comparison in which only the setting varies, institutions choose their training strategies without evidence on this point. In this trial, simulation is the subject of the research and not only its method: the independent variable is the physical setting in which an otherwise identical simulation-based intervention is delivered. The present trial assessed the impact of the physical setting of the simulation on the acquisition of clinical knowledge and on the interprofessional collaborative competencies of perioperative teams, with all the other parameters of the intervention held constant.

Methods

Design, setting and ethical considerations

This was a multicentre, parallel-group, unblinded randomised controlled trial with 1:1 allocation and a single independent variable, the physical setting. In situ simulation was the experimental setting, and a simulation centre configured to replicate the clinical service was the active comparator. The individual professional was the unit of randomisation and of analysis. The intervention, however, was delivered to teams of seven constituted after allocation, so participants are clustered within teams inside each setting, a structure described as an individually randomised group-treatment trial. The framework was exploratory: no outcome was designated as primary, no equivalence or non-inferiority margin was pre-specified, and the two settings were compared with two-sided tests.

The trial was conducted at three locations in Portugal. The experimental setting comprised two perioperative services of different institutions, the Operating Theatre of Unidade Local de Saúde de Matosinhos, Hospital Pedro Hispano (BO-ULSM), and the Post-Anaesthesia Care Unit of Unidade Local de Saúde de São João (UPA-ULSSJ), both of high interprofessional interdependence and of decision-making under time pressure. The active comparator was the Simulation Centre of the Faculty of Medicine of the University of Porto (SIMFMUP), which took no part in recruitment or allocation. Of the 98 participants, 49 were assigned to the in situ setting and 49 to the simulation centre. No participant took part more than once.

The clinical scenario involved local anaesthetic systemic toxicity (LAST), an anaesthetic event of low incidence that combines a response requiring technical skills and a strict sequence of pharmacological interventions with demanding non-technical behaviour on the part of the team, and for which simulation is practically the only means of training [17].

The study complied with the Declaration of Helsinki, Good Clinical Practice and the General Data Protection Regulation, and obtained a favourable opinion from the ethics committees of the Faculty of Medicine of the University of Porto and RISE-Health, of ULSM and of ULSSJ, whose approval numbers are given in the Declarations. Patients and the public were not involved in the design, conduct, reporting or dissemination of this educational research, and no patient data were used.

Participants and team composition

The target population comprised the professionals who respond to an anaesthetic emergency in these services: anaesthetists, perioperative nurses and healthcare assistants (HCAs). In Portugal, healthcare assistants are a regulated, non-clinical professional category who work under medical and nursing supervision and carry out a substantial part of the operation of the perioperative circuit.

Professionals were included if they had at least three months of practice in the perioperative setting of one of the services, were available for the assigned session and gave written informed consent. Pregnant professionals and those who could not complete all phases of the study were excluded.

Each team comprised seven members, two anaesthetists, four nurses and one healthcare assistant, thereby replicating the real response to an anaesthetic emergency, in which the core team of one anaesthetist, three nurses and one healthcare assistant is reinforced during a crisis by a second anaesthetist and a supporting nurse. Fourteen teams were formed, seven per setting, giving a total of 98 participants.

Recruitment and psychological safety

Recruitment was voluntary, non-probabilistic and by convenience, since the aim was not a sample representative of a general population but teams with the real functional composition of an anaesthetic emergency. In each institution it followed the same sequence: an institutional meeting with service directors, an in-person presentation to the teams, a formal invitation by e-mail, self-enrolment on an electronic form and communication of the allocation and the dates. No professional was designated by the service, and all could withdraw at any time without giving a reason. The in situ sessions at BO-ULSM and the sessions at the simulation centre were held between 3 March and 26 March 2026, and the UPA-ULSSJ sessions between 16 and 23 April 2026.

The presentation sessions served a purpose beyond dissemination and formed part of the study's psychological safety strategy. They involved the professionals in the methodology, gave them the opportunity to ask questions and made clear, from the first contact, that the intervention was educational and not assessed. A period for reflection always intervened between the invitation and the decision.

A confidentiality agreement was established with the participants, who undertook not to disclose the scenario or the content of the debriefing to colleagues who had not yet taken part, to prevent contamination between settings.

Randomisation and allocation concealment

Participants were randomised individually, in a 1:1 ratio, to the in situ setting or to the simulation centre. The allocation sequence was computer-generated by an investigator who took no part in recruitment or data collection. Allocation was restricted by blocks defined by professional category within each institution, so that equal numbers of anaesthetists, nurses and healthcare assistants from each institution were allocated to each setting. This restriction was needed because the intervention had to be delivered to complete teams and both institutions had to be represented in both settings. The blocks therefore corresponded to the professional category within each institution and not to permuted blocks of a predetermined size.

The same independent investigator applied the allocation, which remained concealed until the teams were formally constituted; the principal investigator acted only in logistical coordination and accessed the allocations after the sequence had been closed. No blinding was implemented at any stage. The setting could not be concealed from participants and facilitators, the outcomes were self-reported or scored automatically, so there were no separate outcome assessors, and the analyst was not blinded. The analysis used files anonymised through alphanumeric codes that the participants generated themselves and that only they know, so no key linking a code to a named professional was ever held by the research team. Besides allocation concealment, bias was limited by a scenario and a debriefing script fixed in advance and identical in both settings, and by the automatic scoring of the knowledge questionnaire.

Equivalence matrix of the settings

Comparing settings requires that no other element of the intervention varies with the location. Equivalence was ensured by a standardisation matrix in which eleven operational parameters, from the scenario and the simulator to the facilitation team and the timing of the phases, were held constant and only the physical setting varied (Table 1). The same medium-fidelity manikin, control system and monitor, identified in Table 1, were used in every session.

Table 1
Equivalence matrix of the physical simulation settings

Operational parameters	Same	Different	Specification
LAST clinical scenario	Yes		Same script, programmed vital signs and objectives
Medium-fidelity simulator	Yes		Resusci Anne (Laerdal Medical, Stavanger, Norway) operated by SimPad PLUS with SimSTART monitor
Clinical equipment and drugs	Yes		Real equipment in both settings, with identical layout in the two contexts
Learning objectives	Yes		Identical technical and non-technical objectives
Duration of the scenario	Yes		20 minutes
Model and duration of the debriefing	Yes		PEARLS, 40 minutes, co-debriefing
Structure and timing of the session	Yes		Seven phases, 180 minutes in total
Facilitation team	Yes		Two constant facilitators
Technical team	Yes		One simulation technician, responsible for operating the simulator and the vital signs programmed for the progression of the scenario.
Data collection instruments	Yes		Same set before and after the simulation
Spatial layout	Yes		Centre configured to replicate the in situ service
Physical setting		Yes	In situ (BO-ULSM and UPA-ULSSJ) versus centre (SIMFMUP): independent variable

The prior configuration of the simulation centre rooms is a fundamental part of the comparison. Without it, the trial would set a real service against a generic space, and any difference observed could be attributed to architectural familiarity rather than to the setting as such. The centre was therefore organised to replicate the layout of the workplaces (Fig. 1).

The in situ sessions took place in services that remained in operation. Clinical activity continued in the rest of the unit, with the noise that this entails, and the participants' line managers were at work in the service, although none of them attended the sessions. Participants in situ also trained in a space they knew well. These conditions were not controlled, because they are part of what distinguishes training in one's own workplace from training in a centre.

Two pilot sessions, one in each setting, were held beforehand with professionals outside the study cohort to validate the scenario triggers, the timing of the phases and the electronic data capture. No

data from these sessions were analysed, and none of the adjustments affected the scenario, the objectives, the timing or the instruments.

Clinical scenario and participant pathway

The scenario was developed specifically for this trial from the institutional procedures of the two services and from the 2020 ASRA checklist on LAST [17], following the INACSL standards for simulation design [18]. The technical objectives focused on early identification of the neurological and cardiovascular features, on the structured call for help, on the priority administration of 20% lipid emulsion and on avoiding contraindicated agents, progressing to advanced airway management and advanced life support in a shockable rhythm (Additional File 1).

The script included triggers designed to require the active participation of every team member, including the healthcare assistants, whose tasks placed them within the critical path of the response [19]. The initial clinical context was given only to the five professionals of the first-line team, which preserved the realism of the call for help.

Each session lasted 180 minutes, in seven phases of fixed duration always carried out in the same order (Additional File 2). Standardised timing was a requirement of the design, since variation in the duration of the phases would be indistinguishable from an effect of the setting.

In the pre-briefing the learning objectives were made explicit, and the fiction contract was established. The facilitators also set the conditions of psychological safety: they stated that the session was educational and not assessed, that nothing observed or said would be reported to the participants' line managers or institutions, and that participants were free to stop at any moment [20]. The environment and scenario briefings then covered the setting, the equipment, the simulator and the allocation of the real professional roles.

Facilitation and debriefing

All sessions, in both settings and in both in situ services, were conducted by the same core team of three: the principal investigator, a nurse and simulation instructor, acting as co-facilitator; an anaesthetist, also a simulation instructor and co-facilitator; and a specialist simulation technician. The two facilitators were the same in every session. Both hold simulation instructor training, and each had four years of experience in interprofessional clinical simulation.

The debriefing followed the PEARLS model [21] with a standardised script [22] (Additional File 3), as interprofessional co-debriefing structured around the facilitators' observation grid (Additional File 4). PEARLS was chosen because it brings self-assessment, focused facilitation and directive feedback together in a framework that can be standardised across settings. The interprofessional composition of the pair was deliberate, since co-facilitators act as role models for the collaboration the intervention

seeks to develop [23]. In line with the INACSL Sim-IPE standards [24], the script included a phase devoted to the operational role of healthcare assistants in managing the crisis.

Instruments and variables under study

The instruments served specific and coordinated functions in the study design, with no pre-specification of a single primary outcome and no prior hierarchy between them, all results therefore being exploratory in nature. All participants completed the same instruments, administered to the whole team in the same room immediately before and immediately after the session. Table 2 sets out each instrument, its scoring and the timing of administration.

Table 2
Instruments, constructs, scoring and timing of administration

Instrument	Construct	Items and score range	Administration	Analysis metric
Knowledge questionnaire (purpose-built)	Clinical knowledge of LAST	15 items; 0 to 15 correct answers, also reported as percentage	Before and after the session	Within-participant change
RIPLS [25]	Readiness for interprofessional learning	19 items, 5-point Likert; 19 to 95 across four subscales; items 10 to 12 reverse scored ($Y = 6 - X$)	Before the session only	Baseline characteristic
ICCAS [28, 29]	Interprofessional collaborative competencies	20 items, 5-point Likert; 20 to 100 across six domains	Before and after the session	Within-participant change
SET-M [31, 32]	Perceived educational effectiveness	19 items after exclusion of the observer item, 3-point scale; 19 to 57	After the session	Score at that time point
SDS [33]	Perceived quality of the scenario design	20 items, 5-point Likert; two parallel subscales of 20 to 100 (presence and importance)	After the session	Score at that time point
DASH, participant version, short form [34]	Structural quality of the debriefing	6 elements scored 1 to 7; total up to 42	Immediately after the debriefing	Score at that time point

The Readiness for Interprofessional Learning Scale (RIPLS) was administered before the simulation, in the four-subscale version [25], with reverse scoring of the Negative Professional Identity subscale. Its

purpose was to characterise the baseline comparability of the teams in terms of readiness for interprofessional learning; it was not used as a measure of change, given the scale's well-documented ceiling effect [26].

The multidimensional assessment of learning comprised two dimensions: clinical knowledge of LAST, assessed with a 15-item multiple-choice questionnaire, developed for the study according to Bloom's taxonomy and refined after a focus group with experts [27] (Additional File 5). The items of the knowledge questionnaire were written to be answerable across professions, and for healthcare assistants they address recognition of the event and of their own role in the response rather than clinical management. Interprofessional collaborative competencies, measured with the Interprofessional Collaborative Competencies Attainment Scale (ICCAS), a 20-item instrument covering six domains [28, 29]. The ICCAS was administered directly before and after the simulation, a methodological adaptation of the original retrospective format intended to mitigate recall and social desirability bias [30].

Reaction and perceived educational effectiveness were measured after the session with the Simulation Effectiveness Tool-Modified (SET-M) [31, 32], in 19 items after exclusion of the item on the observer role, since every participant took an active part in the scenario.

The Simulation Design Scale (SDS) [33] and the short form of the Debriefing Assessment for Simulation in Healthcare (DASH, participant version) [34] served as psychometric tools for auditing the process. They did not measure learning, but assessed whether the architecture of the scenario (fidelity, objectives and support) and the facilitation of the debriefing had been delivered in a structurally identical way in the two study contexts, in situ and in the simulation centre, supporting the equivalence matrix.

None of these instruments has a cross-cultural adaptation validated for European Portuguese. The RIPLS, the ICCAS and the SDS were administered in the original English version, the SET-M in the version translated, adapted and validated for Brazilian Portuguese by Bergamasco and Cruz [32], obtained from the healthcare simulation assessment instruments page of HealthySimulation.com [35], and the DASH in the Portuguese translation released by the Center for Medical Simulation. All data were collected electronically, and the responses given before and after the session were paired by an alphanumeric code that each participant generated and that only that participant knows. All outcomes were measured within the session, and there was no follow-up.

Sample size and statistical analysis

The sample size was calculated a priori in G*Power 3.1.9.7 [36] for a two-sided independent-samples t test, assuming an effect size of $d = 0.60$, a significance level of 0.05 and statistical power of 0.80. Because no outcome was designated as primary, the calculation refers to a generic comparison between two groups. It required 45 participants per setting, and the sample was set at 49 per setting so that complete teams of seven could be formed.

Statistical analysis was carried out in IBM SPSS Statistics, version 30, with a significance level of 0.05. Continuous variables are presented as mean $\pm$ standard deviation, to allow comparison with prior literature and because the distributions were approximately symmetric, and categorical variables as frequencies and percentages. Non-parametric tests were chosen as a conservative approach for bounded ordinal scales. Comparisons between settings used the Mann-Whitney U test, and Pearson's chi-squared test for sex; changes within each setting were analysed with the Wilcoxon signed-rank test.

The analysis was computed on individual responses, and also by professional category. All 98 participants were analysed in the setting to which they had been allocated, and within their professional category, with no crossovers and no exclusions, and there were no missing data, since the electronic forms required an answer to every item.

Changes to the trial protocol

The changes to the protocol were decided before the first participant was enrolled, and no outcome data were examined before the end of data collection. Additional File 6 lists the nine changes with their reasons and dates. Three of them bear on interpretation: a second in situ site was added, in order to reach the planned sample within the available period and to include teams from more than one institution; the assumed effect size was revised from 0.50 to 0.60, which set the sample at 98 participants; and detection of latent safety threats, listed among the protocol outcomes, was not an objective of the trial as conducted and no validated instrument was available, so it was not measured.

We used the CONSORT 2025 reporting guideline [37] to draft this manuscript and the CONSORT reporting checklist [38] when editing it (Additional File 7), together with the reporting guidelines for simulation-based research of Cheng and colleagues [39]. The simulation-based intervention is reported against the simulation-specific items of those guidelines in Additional File 8.

Results

Participant flow and baseline comparability

Throughout this section, and in every table, the value for the in situ setting is given first and the value for the simulation centre second. All 98 professionals assessed for eligibility met the inclusion criteria, consented to take part and were randomised 1:1 to in situ simulation ($n = 49$) or to the simulation centre ($n = 49$), and were then organised into 14 teams of seven (Fig. 2). The sample comprised 28 anaesthetists (28.6%), 56 perioperative nurses (57.1%) and 14 healthcare assistants (14.3%). All participants completed the session and all measurements, with no losses and no exclusions. Recruitment ended when the planned number of complete teams had been formed, and no adverse event was recorded in either setting.

The two settings did not differ on any of the characterisation variables (Table 3). Mean age was 44.14 ± 10.41 years in situ and 43.16 ± 10.23 in the centre ($p = 0.508$), professional experience was 20.00 ± 10.32

and 18.14 ± 10.85 years ($p = 0.380$), and the proportion of women was 89.8% and 79.6% ($p = 0.161$). Previous experience with clinical simulation was reported by 49.0% and 46.9% of participants ($p = 0.841$). Baseline RIPLS scores did not differ on any subscale or on the total (81.73 ± 6.67 and 81.04 ± 9.76 ; $p = 0.809$).

Table 3

Baseline characterisation of the sample by study group (physical setting): sociodemographic and professional data and readiness for interprofessional learning (RIPLS, pre-simulation)

	In Situ (n = 49)	Centre (n = 49)	p-value*
Sociodemographic and professional data			
Age (years)	44.14 ± 10.41	43.16 ± 10.23	0.508*
Sex, n (%)	M 5 (10.2); F 44 (89.8)	M 10 (20.4); F 39 (79.6)	0.161**
Professional experience (years)	20.00 ± 10.32	18.14 ± 10.85	0.380*
Previous simulation experience (%)	48.98 ± 50.51	46.94 ± 50.42	0.841**
RIPLS			
Teamwork and collaboration [1–9]	41.90 ± 3.44	41.18 ± 5.42	0.708*
Negative Professional Identity [10–12], reversed	13.04 ± 2.71	13.02 ± 2.67	0.831*
Positive Professional Identity [13–16]	17.27 ± 2.12	17.70 ± 2.76	0.160*
Roles and Responsibilities [17–19]	9.53 ± 1.96	9.14 ± 1.88	0.298*
Total	81.73 ± 6.67	81.04 ± 9.76	0.809*
Note: age, professional experience and RIPLS scores expressed as mean $\pm$ SD (standard deviation); gender expressed as absolute and relative frequency.			
*Non-parametric test: Mann-Whitney U. A p-value < 0.05 is statistically significant.			
**Pearson's chi-squared test for categorical data (sex). A p-value < 0.05 is statistically significant.			

Equivalence of the intervention between the two settings

Before comparing learning, we verified whether the intervention had been delivered in the same way in the two settings (Table 4). No domain of the SDS differed, on either the design-elements subscale ($p = 0.613$) or the importance subscale ($p = 0.182$). Fidelity deserves separate mention, because it is the domain in which unequal delivery would show most directly: participants who trained in their own service rated it at the same level as those who trained in the centre (9.43 ± 1.12 and 9.31 ± 0.96 ; $p = 0.305$). Perceived effectiveness on the SET-M was also comparable (51.18 ± 5.16 and 51.71 ± 5.77 ; $p = 0.366$).

Table 4

Perceived quality of the scenario design (SDS), perceived effectiveness (SET-M) and structural quality of the debriefing (DASH - participant version) by physical setting (post-simulation)

SDS				
	Subscale	In Situ (n = 49)	Centre (n = 49)	p-value*
Objectives and Information [1–5]	Simulation (General)	23.49 ± 2.38	23.39 ± 2.15	0.878
	For You (Individual)	23.14 ± 2.72	23.35 ± 2.23	0.757
Support [6–9]	Simulation (General)	17.98 ± 2.67	18.69 ± 1.99	0.089
	For You (Individual)	17.98 ± 2.17	18.61 ± 1.64	0.181
Problem Solving [10–14]	Simulation (General)	22.29 ± 2.63	22.45 ± 2.38	0.885
	For You (Individual)	21.82 ± 2.60	22.37 ± 2.60	0.254
Feedback/Guided Reflection [15–18]	Simulation (General)	19.20 ± 1.32	19.18 ± 1.64	0.713
	For You (Individual)	18.84 ± 1.50	19.18 ± 1.39	0.158
Fidelity [19–20]	Simulation (General)	9.43 ± 1.12	9.31 ± 0.96	0.305
	For You (Individual)	9.37 ± 1.15	9.51 ± 0.79	0.752
Total	Simulation (General)	92.39 ± 8.18	93.02 ± 7.00	0.613
	For You (Individual)	91.14 ± 8.45	93.02 ± 7.26	0.182
Simulation Effectiveness Tool-Modified (SET-M)				
Note: values expressed as mean ± SD.				
*Non-parametric test: Mann-Whitney U. A p-value < 0.05 is statistically significant.				

SDS				
	Subscale	In Situ (n = 49)	Centre (n = 49)	p-value*
		In Situ (n = 49)	Centre (n = 49)	p-value*
Sub-Scales				
	Pre-briefing	5.55 ± 0.79	5.27 ± 1.08	0.183
	Learning	15.55 ± 2.11	15.92 ± 2.51	0.167
	Confidence	15.39 ± 2.98	15.98 ± 2.19	0.549
	Debriefing	14.69 ± 0.85	14.55 ± 1.10	0.850
	Total	51.18 ± 5.16	51.71 ± 5.77	0.366
Debriefing Assessment for Simulation in Healthcare (DASH), participant version				
		In Situ (n = 49)	Centre (n = 49)	p-value*
	Element 1 - Introduction	6.57 ± 0.68	6.69 ± 0.58	0.351
	Element 2 - Learning environment	6.59 ± 0.67	6.71 ± 0.54	0.421
	Element 3 - Structure of the debriefing	6.57 ± 0.61	6.65 ± 0.60	0.426
	Element 4 - Reflection on performance	6.37 ± 0.76	6.67 ± 0.55	0.036
	Element 5 - Strengths and weaknesses	6.57 ± 0.65	6.51 ± 0.68	0.612
	Element 6 - Performance improvement	6.65 ± 0.60	6.65 ± 0.56	0.887
	Total	39.33 ± 3.56	39.90 ± 2.76	0.677
Note: values expressed as mean ± SD.				
*Non-parametric test: Mann-Whitney U. A p-value < 0.05 is statistically significant.				

The DASH totals were 39.33 ± 3.56 in situ and 39.90 ± 2.76 in the centre (p = 0.677), and five of the six elements did not differ. The exception was element 4, on how far the debriefing helped participants to analyse their own performance, which was rated lower in situ (6.37 ± 0.76) than in the centre (6.67 ± 0.55; p = 0.036). This was the only statistically significant result in the study. With this exception, the participants' own assessment confirms that the scenario and the facilitation were equivalent, which supports attributing to the physical setting any difference found in the learning measures.

Learning and participants' reaction

The ICCAS rose from 70.59 ± 13.80 to 80.45 ± 13.33 in the in situ group and from 68.83 ± 10.72 to 78.10 ± 12.49 in the centre group, with significant improvement in both ($p < 0.001$) and no difference between settings before ($p = 0.614$) or after the session ($p = 0.335$) (Table 5). The pattern held across the six domains, all of which improved within each setting and none of which differed between settings (post-simulation p values between 0.141 and 0.756).

Table 5
Interprofessional collaborative competencies (ICCAS) by physical setting (pre- and post-simulation)

		In Situ (n = 49)	Centre (n = 49)	p-value*
Communication [1–5]	Pre	18.10 ± 4.08	17.31 ± 3.09	0.383
	Post	20.14 ± 3.98	19.92 ± 3.38	0.594
	p**	< 0.001	< 0.001	
Collaboration [6–8]	Pre	10.67 ± 2.32	10.41 ± 1.77	0.786
	Post	12.00 ± 2.14	11.96 ± 1.81	0.756
	p**	< 0.001	< 0.001	
Roles and Responsibilities [9–12]	Pre	13.69 ± 2.96	13.76 ± 2.27	0.894
	Post	16.37 ± 2.47	15.82 ± 2.70	0.319
	p**	< 0.001	< 0.001	
Patient-Centered Care [13–15]	Pre	9.90 ± 2.19	10.12 ± 1.91	0.476
	Post	11.39 ± 2.11	10.96 ± 2.12	0.276
	p**	< 0.001	0.002	
Conflict Management [16–18]	Pre	11.18 ± 1.91	10.59 ± 1.95	0.128
	Post	12.45 ± 1.87	11.80 ± 2.18	0.149
	p**	< 0.001	< 0.001	
Team Functioning [19–20]	Pre	7.04 ± 1.68	6.65 ± 1.22	0.114
	Post	8.10 ± 1.61	7.65 ± 1.69	0.141
	p**	< 0.001	< 0.001	
Total	Pre	70.59 ± 13.80	68.83 ± 10.72	0.614
	Post	80.45 ± 13.33	78.10 ± 12.49	0.335
	p**	< 0.001	< 0.001	
Note: values expressed as mean ± SD.				
*Non-parametric test: Mann-Whitney U. A p-value < 0.05 is statistically significant.				
**Non-parametric test: Wilcoxon Signed Ranks. A p-value < 0.05 is statistically significant.				

Knowledge, expressed as the percentage of correct answers out of a maximum of 100%, was 77.82 ± 25.14% in the in situ group and 82.18 ± 22.58% in the centre before the session (p = 0.253). After the

intervention it rose to $87.35 \pm 19.26\%$ and $90.48 \pm 19.77\%$, with a significant increase in both ($p < 0.001$) and no difference between settings ($p = 0.103$) (Table 6). Across the learning measures, therefore, the data show no evidence of educational loss in either context.

Table 6
Change in knowledge of LAST by physical setting (pre- and post-simulation)

Knowledge (LAST)	In Situ (n = 49)	Center (n = 49)	p-value*
Pre-simulation (%)	77.82 ± 25.14	82.18 ± 22.58	0.253
Post-simulation (%)	87.35 ± 19.26	90.48 ± 19.77	0.103
(Change) p-value**	< 0.001	< 0.001	
Note: values expressed as mean $\pm$ SD.			
*Non-parametric test: Mann-Whitney U. A p-value < 0.05 is statistically significant.			
**Non-parametric test: Wilcoxon Signed Ranks. A p-value < 0.05 is statistically significant.			

Discussion

With the eleven operational parameters of the intervention held strictly identical through the equivalence matrix (Table 1), no statistically significant difference was detected between in situ simulation and the dedicated centre on any of the measures of clinical learning, attitudes or participant reaction. Clinical knowledge of LAST and interprofessional collaborative competencies improved markedly in both settings and to a similar extent (Tables 5 and 6).

These results allow the competing hypotheses in the literature to be weighed. The expectation that in situ training would be superior rests on the argument that the fidelity and authenticity of the environment favour transfer to clinical practice [13], but it had never been tested in a comparison in which only the setting varied. The data do not confirm it, and the most direct argument comes from the participants themselves, who rated the fidelity of the session at the same level in the centre as in their own service.

This finding is consistent with the position of Norman and colleagues, according to which environmental fidelity plays a secondary role in the transfer of skills when compared with the functional alignment of the task and the quality of facilitation [14, 15]. Provided that the clinical stimuli, the monitoring equipment and the care tasks are authentic, the data do not suggest that the surrounding physical space adds to the effectiveness of learning.

The absence of educational asymmetry is consistent with the evidence obtained in obstetric emergencies by Crofts and colleagues and by Sørensen and colleagues [11, 12]. The contribution of this trial lies in the degree of control under which the result was obtained in the perioperative setting, and in the fact that the standardisation of the intervention was not merely assumed in the protocol but verified in the participants' own data (Table 1).

The two settings were compared with two-sided tests, and no statistically significant difference was identified between them. Since the trial was exploratory, with no equivalence margin pre-specified and a sample size that was not calculated to demonstrate one, the study does not claim equivalence in the technical sense of the term.

What the trial establishes is that there is no evidence of educational loss between the two contexts, that is, none associated with training in one's own service rather than in a dedicated centre, under equivalent conditions of design and facilitation. It is on this absence of evidence of harm, and not on any claim of statistical equivalence, that the institutional decisions discussed below rest.

Clinical governance, frequency and institutional sustainability

With no evidence of educational harm in local training, hospitals no longer depend exclusively on dedicated centres to deliver high-quality team training. This is relevant in an international context of unequal access to specialist centres [5].

Logistical sustainability and the frequency of training then become the determining variables. Pisciotanni and colleagues found that in situ simulation repeated at four-month intervals sustained knowledge and technical performance, whereas the group without periodic practice deteriorated [40]. In situ simulation removes the need for team travel and prolonged clinical leave, so training can become more frequent without additional off-service hours, and team training may become a structural part of hospital clinical governance, as called for by the 2024 Global Consensus Statement and the 2026 ESAIC Østergaard Declaration [3, 4].

Equity of access and representation of the complete clinical team

Delivering training in the workplace also increases access to it. Training in a dedicated centre requires releasing professionals from the service, and when resources are limited institutions tend to release those whose absence is easiest to justify, which favours doctors and nurses and leaves out professionals with supporting roles in care delivery. By drawing on the staff on duty and integrating healthcare assistants as full members of the response to the LAST crisis, this trial ensured the ecological validity of the scenario. It also showed that the team can be trained in its real composition, with hierarchy levelled in the debriefing, which meets the aim of educational equity set out in the INACSL Sim-IPE standards [24].

Translational simulation and optimisation of perioperative workflows

Although this trial assessed outcomes at the levels of reaction and learning (Kirkpatrick levels 1 and 2), the literature on translational simulation stresses that in situ simulation offers a functional diagnosis of

the real system that a dedicated centre does not reproduce [6, 7, 41]. Training in the workplace exposes actual distances, the visibility of algorithms, the time taken to obtain drugs and emergency devices, and the telephone pathways for calling for help. Since the present study found no evidence of educational loss in situ, the potential to identify latent safety threats becomes an additional gain with no known educational trade-off.

Critical reflection and psychological safety in the workplace

The only statistically divergent result was on element 4 of the DASH, relating to the depth with which facilitators helped participants to explore their clinical reasoning during the debriefing, rated lower in the hospital than in the simulation centre (6.37 ± 0.76 vs. 6.67 ± 0.55 ; $p = 0.036$; Table 4).

This result corresponds closely with the literature on human factors and psychological safety. As formulated by Edmondson and operationalised by Rudolph and colleagues in the concept of the safe container established in the pre-briefing, honest analysis of error depends on the perception of safety for interpersonal risk [42, 43].

Krarpur and colleagues describe in situ simulation with authentic roles as having a dual nature: while it maximises sociological realism, it also increases the learner's exposure before colleagues and direct line managers, raising the perceived risk to professional image and reputation [16, 44]. The simulation centre, by contrast, offers a neutral space, physically removed from workplace pressures.

The conditions of the in situ sessions in this trial fit that description. Care continued in the service while the sessions ran, with the noise of a unit in operation, and the participants' line managers were at work nearby, even though none of them attended. Noise may have reduced the concentration available for reflection, and the proximity of line managers may have made participants more guarded in analysing their own performance, despite the confidentiality commitments made in the pre-briefing. Neither factor was present in the centre, which offers a plausible explanation for the lower rating of element 4 in situ.

This finding supports a direct operational recommendation: in situ simulation demands redoubled investment by facilitators in the pre-briefing, establishing the fiction contract and the rules of confidentiality clearly and dismantling the fear of punitive judgement before the scenario begins. Where the service allows it, holding the debriefing in a room protected from the activity of the unit may also help. As an isolated finding without adjustment for multiplicity, it should be read as hypothesis-generating.

Limitations

Randomisation was individual, but participants trained in teams and shared the scenario, the facilitators and the debriefing. The tests treated their data as independent, so the genuinely independent information lies somewhere between the 98 professionals and the 14 teams. The p values are therefore an optimistic estimate of significance, and differences close to the conventional threshold, such as that

on element 4 of the DASH, call for added caution. With seven teams per setting, the number of clusters did not allow this effect to be corrected by mixed models with adequate precision.

The trial was not designed as a formal equivalence study with pre-established margins, so the absence of significant differences reflects a lack of evidence of an effect and not a demonstration of equivalence. The use of several instruments, without a primary outcome or correction for multiple testing, also makes all the results exploratory.

None of the scales has been validated for European Portuguese. We chose not to translate them, preserving their original structure, but responding in a language that is not one's native language may cause hesitation and introduce noise into the responses. The ICCAS was also administered outside the retrospective format in which it was validated. A further limitation concerns the populations in which these instruments were developed: the RIPLS, the ICCAS, the SET-M and the SDS were validated in samples of doctors and nurses, or of students in those professions, and none of them was designed for professional categories with supporting roles in care delivery. Applying them to healthcare assistants therefore rests on an assumption of measurement equivalence that has not been tested, and instruments built for the complete functional team, including these professionals, are still needed.

The convenience sample and the open-label design, in two services in the same country, limit generalisability. The data are confined to the first two Kirkpatrick levels, with no follow-up, and rest on self-report and a knowledge test after a single scenario and a single exposure, with no observed measure of team performance. The simulation centre was configured to replicate the workplace, so the findings apply to centres with equivalent structural fidelity and not to generic spaces. Noise and the presence of line managers in the in situ services were not measured, so their contribution to the DASH result cannot be quantified.

The knowledge questionnaire showed a floor effect among healthcare assistants and a ceiling effect among anaesthetists, which limits the capacity of these scores to detect change. No reliability coefficients were computed in this sample. Finally, because the number of participants in each stratum was fixed in advance, the allocation of the last participants of each professional category became predictable, although it remained concealed until the teams were constituted.

Future research

Future research should proceed in the translational direction, with non-inferiority trials using a margin defined a priori and a sample size calculated for it, and with longitudinal studies able to demonstrate effects on clinical performance and patient outcomes.

Conclusions

In this randomised controlled trial conducted in the perioperative setting, with strict equivalence of intervention and facilitation, interprofessional teams trained in their own clinical service and those

trained in a dedicated centre showed comparable gains in clinical knowledge and in interprofessional collaborative competencies. The only asymmetry observed, favouring the centre in the depth of reflection on performance during the debriefing, points to the need for a structured safe container when training in the workplace.

The absence of an educational disadvantage of in situ simulation removes support from the premise that effective team training depends on dedicated simulation centres, so institutional decisions may be guided by accessibility and sustainability. Because the complete functional team can train without travelling and with less demand on staff release, training in the workplace makes regular cycles of training possible and allows simulation to be used for the continuous identification of latent threats.

Use of large language models

Generative artificial intelligence tools were used to support language editing and the structural organisation of the manuscript. Additionally, while the supplementary materials (Additional Files 1, 2, 3, 4, and 5) were originally developed and administered in Portuguese in the clinical setting, their translation into English was performed using Claude (Anthropic). The tools were not used to generate original data, conduct statistical analyses, or formulate scientific conclusions. All authors critically reviewed, verified, and approved the translated materials and the final version of the manuscript, taking full responsibility for all content.

Abbreviations

ASRA

American Society of Regional Anesthesia and Pain Medicine

BO-ULSM

Operating Theatre, Unidade Local de Saúde de Matosinhos

CONSORT

Consolidated Standards of Reporting Trials

DASH

Debriefing Assessment for Simulation in Healthcare

ICCAS

Interprofessional Collaborative Competencies Attainment Scale

INACSL

International Nursing Association for Clinical Simulation and Learning

LAST

local anaesthetic systemic toxicity

PEARLS

Promoting Excellence and Reflective Learning in Simulation

RIPLS

Readiness for Interprofessional Learning Scale

SDS
Simulation Design Scale
SET-M
Simulation Effectiveness Tool-Modified
SIMFMUP
Simulation Centre, Faculty of Medicine, University of Porto
Sim-IPE
Simulation-Enhanced Interprofessional Education
HCA
healthcare assistant
UPA-ULSSJ
Post-Anaesthesia Care Unit, Unidade Local de Saúde de São João

Declarations

Ethics approval and consent to participate:

Ethics approval and consent to participate: Approved by the Ethics Committee of the Faculty of Medicine of the University of Porto and RISE-Health (393/CEFMUP-RISEHealth/2025), by the Health Ethics Committee of ULSM, with authorisation from its board (process 20250125), and by the Health Ethics Committee of ULSSJ (process 53/2026). The study involved no patients and no patient data, and the in situ sessions did not interfere with patient care. All participants gave written, free and informed consent.

Consent for publication:

Written informed consent for publication of identifiable clinical simulation images was obtained from all healthcare professionals visible in Fig. 1.

Competing interests:

MCC, JJS, RR and AN declare that they have no competing interests. No author has any financial or intellectual property interest in the simulator, in its control system, or in any of the assessment instruments used in this trial.

Funding:

This work was supported by National Funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., within the project RISE-Health - UID/06397/2025.

Author Contribution

All authors have made substantial contributions to the manuscript, namely:• MCC conceived and designed the study, obtained the ethical approvals, prepared the study documents, recruited the participants, co-facilitated the sessions and debriefings, collected and analysed the data, and drafted the manuscript;• JJS co-supervised the study, contributed to the development of the clinical scenario, the knowledge questionnaire and the standardised debriefing script, co-facilitated the sessions and debriefings, and reviewed the manuscript;• RR contributed to the development of the clinical scenario and the standardised debriefing script, and reviewed the manuscript;• AN supervised the study, contributed to its conception, methodological design and statistical approach, analysed the data, acted as specialist simulation technician, and reviewed the manuscript;• All authors reviewed it critically;• All authors have approved the manuscript version submitted and agree with its submission.

Acknowledgement

The authors thank the perioperative teams of the Operating Theatre of Unidade Local de Saúde de Matosinhos and of the Post-Anaesthesia Care Unit of Unidade Local de Saúde de São João, and the staff of the Simulation Centre of the Faculty of Medicine of the University of Porto, Portugal.

Data Availability

The data generated and analysed in this study are not publicly available. Participants consented to statistical treatment and to dissemination of the results exclusively in aggregated and anonymised form, by professional group, so individual participant data cannot be shared. Aggregated data underlying the results reported here are available from the corresponding author on reasonable request. Records are held on a secure institutional server of the Faculty of Medicine of the University of Porto, under restricted access, and are retained for a maximum of one year after the end of the study, as approved by the research ethics committee, after which they are permanently and securely deleted. The knowledge questionnaire, the clinical scenario, the debriefing script, the observation grid and the session plan are available as additional files. The study protocol is available with the trial registration record at <https://clinicaltrials.gov/study/NCT07346937>.

References

1. Weaver SJ, Rosen MA, DiazGranados D, Lazzara EH, Lyons R, Salas E, et al. Does teamwork improve performance in the operating room? A multilevel evaluation. *Jt Comm J Qual Patient Saf*. 2010;36(3):133–42. [http://doi.org/10.1016/s1553-7250\(10\)36022-3](http://doi.org/10.1016/s1553-7250(10)36022-3).
2. Institute of Medicine (US) Committee on Quality of Health Care in America. In: Kohn LT, Corrigan JM, Donaldson MS, editors. *To err is human: building a safer health system*. Washington (DC): National Academies Press (US); 2000. <http://doi.org/10.17226/9728>.

3. Diaz-Navarro C, Armstrong R, Charnetski M, Freeman KJ, Koh S, Reedy G, et al. Global consensus statement on simulation-based practice in healthcare. *Adv Simul (Lond)*. 2024;9(1):19. <http://doi.org/10.1186/s41077-024-00288-1>.
4. Berger-Estilita J, Burlacu CL, Ostergaard D, Sinai YB, Preckel B, Crisan I, et al. The Ostergaard Declaration: a European Society of Anaesthesiology and Intensive Care statement on advancing simulation-based education and training. *Eur J Anaesthesiol*. 2026;43(9):719–22. <http://doi.org/10.1097/EJA.0000000000002431>.
5. Abramovich I, Crisan I, Dow O, Morais D, De Hert S, Østergaard D, et al. Simulation-based education in anaesthesiology residency training in Europe: a survey-based cross-sectional study. *Trends Anaesth Crit Care*. 2023;53:101310. <http://doi.org/10.1016/j.tacc.2023.101310>.
6. Patterson MD, Geis GL, Falcone RA, LeMaster T, Wears RL. In situ simulation: detection of safety threats and teamwork training in a high risk emergency department. *BMJ Qual Saf*. 2013;22(6):468–77. <http://doi.org/10.1136/bmjqs-2012-000942>.
7. Grace MA, O'Malley R. Using in situ simulation to identify latent safety threats in emergency medicine: a systematic review. *Simul Healthc*. 2024;19(4):243–53. <http://doi.org/10.1097/SIH.0000000000000748>.
8. Schram A, Paltved C, Christensen KB, Kjaergaard-Andersen G, Jensen HI, Kristensen S. Patient safety culture improves during an in situ simulation intervention: a repeated cross-sectional study. *BMJ Open Qual*. 2021;10(1):e001183. <http://doi.org/10.1136/bmjoq-2020-001183>.
9. Goldshtein D, Krensky C, Doshi S, Perelman VS. In situ simulation and its effects on patient outcomes: a systematic review. *BMJ Simul Technol Enhanc Learn*. 2020;6(1):3–9. <http://doi.org/10.1136/bmjstel-2018-000387>.
10. Martin A, Cross S, Attoe C. The use of in situ simulation in healthcare education: current perspectives. *Adv Med Educ Pract*. 2020;11:893–903. <http://doi.org/10.2147/AMEP.S188258>.
11. Crofts JF, Ellis D, Draycott TJ, Winter C, Hunt LP, Akande VA. Change in knowledge of midwives and obstetricians following obstetric emergency training: a randomised controlled trial of local hospital, simulation centre and teamwork training. *BJOG*. 2007;114(12):1534–41. <http://doi.org/10.1111/j.1471-0528.2007.01493.x>.
12. Sorensen JL, van der Vleuten C, Rosthoj S, Ostergaard D, LeBlanc V, Johansen M, et al. Simulation-based multiprofessional obstetric anaesthesia training conducted in situ versus off-site leads to similar individual and team outcomes: a randomised educational trial. *BMJ Open*. 2015;5(10):e008344. <http://doi.org/10.1136/bmjopen-2015-008344>.
13. Lavoie P, Deschênes MF, Nolin R, Bélisle M, Blanchet Garneau A, Boyer L, et al. Beyond technology: a scoping review of features that promote fidelity and authenticity in simulation-based health professional education. *Clin Simul Nurs*. 2020;42:22–41. <http://doi.org/10.1016/j.ecns.2020.02.001>.
14. Norman G, Dore K, Grierson L. The minimal relationship between simulation fidelity and transfer of learning. *Med Educ*. 2012;46(7):636–47. <http://doi.org/10.1111/j.1365-2923.2012.04243.x>.

15. Mitchell S, Blanchard E, Curran V, Hoadley T, Donoghue A, Lockey A. Effects of simulation fidelity on health care providers on team training: a systematic review. *Simul Healthc*. 2024;19(1S):S50–6. <http://doi.org/10.1097/SIH.0000000000000762>.
16. Krarup KB, McKenzie AC, Al-Masri A, Abildgren L. Authentic roles matter: how in situ simulation influences fidelity, psychological safety, and learning transfer. *Adv Simul (Lond)*. 2026. <http://doi.org/10.1186/s41077-026-00449-4>.
17. Neal JM, Neal EJ, Weinberg GL. American Society of Regional Anesthesia and Pain Medicine Local Anesthetic Systemic Toxicity checklist: 2020 version. *Reg Anesth Pain Med*. 2021;46(1):81–2. <http://doi.org/10.1136/rapm-2020-101986>.
18. Committee INACSLs, Watts PI, McDermott DS, Alinier G, Charnetski M, Ludlow J, et al. Healthcare simulation standards of best practice: simulation design. *Clin Simul Nurs*. 2021;58:14–21. <http://doi.org/10.1016/j.ecns.2021.08.009>.
19. Weller J, Boyd M, Cumin D. Teams, tribes and patient safety: overcoming barriers to effective teamwork in healthcare. *Postgrad Med J*. 2014;90(1061):149–54. <http://doi.org/10.1136/postgradmedj-2012-131168>.
20. Committee INACSLs, McDermott DS, Ludlow J, Horsley E, Meakim C. Healthcare simulation standards of best practice: prebriefing, preparation and briefing. *Clin Simul Nurs*. 2021;58:9–13. <https://doi.org/10.1016/j.ecns.2025.101777>.
21. Eppich W, Cheng A. Promoting Excellence and Reflective Learning in Simulation (PEARLS): development and rationale for a blended approach to health care simulation debriefing. *Simul Healthc*. 2015;10(2):106–15. <http://doi.org/10.1097/SIH.0000000000000072>.
22. Bajaj K, Meguerdichian M, Thoma B, Huang S, Eppich W, Cheng A. The PEARLS healthcare debriefing tool. *Acad Med*. 2018;93(2):336. <http://doi.org/10.1097/ACM.0000000000002035>.
23. Cheng A, Palaganas J, Eppich W, Rudolph J, Robinson T, Grant V. Co-debriefing for simulation-based education: a primer for facilitators. *Simul Healthc*. 2015;10(2):69–75. <http://doi.org/10.1097/SIH.0000000000000077>.
24. Committee INACSLs, Rossler K, Molloy MA, Pastva AM, Brown M, Xavier N. Healthcare simulation standards of best practice: simulation-enhanced interprofessional education. *Clin Simul Nurs*. 2021;58:49–53. <http://doi.org/10.1016/j.ecns.2021.08.015>.
25. McFadyen AK, Webster V, Strachan K, Figgins E, Brown H, McKechnie J. The Readiness for Interprofessional Learning Scale: a possible more stable sub-scale model for the original version of RIPLS. *J Interprof Care*. 2005;19(6):595–603. <http://doi.org/10.1080/13561820500430157>.
26. Torsvik M, Johnsen HC, Lillebo B, Reinaas LO, Vaag JR. Has the ceiling rendered the Readiness for Interprofessional Learning Scale (RIPLS) outdated? *J Multidiscip Healthc*. 2021;14:523–31. <http://doi.org/10.2147/JMDH.S296418>.
27. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front Public Health*. 2018;6:149. <http://doi.org/10.3389/fpubh.2018.00149>.

28. Archibald D, Trumpower D, MacDonald CJ. Validation of the interprofessional collaborative competency attainment survey (ICCAS). *J Interprof Care*. 2014;28(6):553–8. <http://doi.org/10.3109/13561820.2014.917407>.
29. Schmitz CC, Radosovich DM, Jardine P, MacDonald CJ, Trumpower D, Archibald D. The interprofessional collaborative competency attainment survey (ICCAS): a replication validation study. *J Interprof Care*. 2017;31(1):28–34. <http://doi.org/10.1080/13561820.2016.1233096>.
30. Kruger JS, Tona J, Kruger DJ, Jackson JB, Ohtake PJ. Validation of the Interprofessional Collaborative Competency Attainment Survey (ICCAS) retrospective pre-test measures. *J Interprof Care*. 2023;37(5):791–6. <http://doi.org/10.1080/13561820.2023.2169261>.
31. Leighton K, Ravert P, Mudra V, Macintosh C. Updating the Simulation Effectiveness Tool: item modifications and reevaluation of psychometric properties. *Nurs Educ Perspect*. 2015;36(5):317–23. <http://doi.org/10.5480/15-1671>.
32. Bergamasco EC, Cruz DALM. Simulation effectiveness tool modified (SET-M): adaptation and validation for Brazil. *Rev Lat Am Enfermagem*. 2021;29:e3437. <http://doi.org/10.1590/1518-8345.4282.3437>.
33. Jeffries PR, Rizzolo MA. Designing and implementing models for the innovative use of simulation to teach nursing care of ill adults and children: a national, multi-site, multi-method study. New York: National League for Nursing and Laerdal Medical; 2006.
34. Brett-Fleegler M, Rudolph J, Eppich W, Monuteaux M, Fleegler E, Cheng A, et al. Debriefing assessment for simulation in healthcare: development and psychometric properties. *Simul Healthc*. 2012;7(5):288–94. <http://doi.org/10.1097/SIH.0b013e3182620228>.
35. HealthySimulation.com. Simulation Effectiveness Tool-Modified (SET-M) [Internet]. <https://www.healthysimulation.com/tools/evaluating-healthcare-simulation/simulation-effectiveness-tool-modified-set-m/>. Accessed 10 September 2026.
36. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175–91. <http://doi.org/10.3758/bf03193146>.
37. Hopewell S, Chan AW, Collins GS, Hrobjartsson A, Moher D, Schulz KF, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. *BMJ*. 2025;389:e081123. <http://doi.org/10.1136/bmj-2024-081123>.
38. Hopewell S, Chan AW, Collins GS, Hrobjartsson A, Moher D, Schulz KF et al. The CONSORT reporting checklist. In: Harwood J, Albury C, de Beyer J, Schlüssel M, Collins G, editors. The EQUATOR Network reporting guideline platform. UK EQUATOR Centre; 2025. <https://resources.equator-network.org/reporting-guidelines/consort/consort-checklist.docx>. Accessed 21 September 2026.
39. Cheng A, Kessler D, Mackinnon R, Chang TP, Nadkarni VM, Hunt EA, et al. Reporting guidelines for health care simulation research: extensions to the CONSORT and STROBE statements. *Adv Simul (Lond)*. 2016;1:25. <http://doi.org/10.1186/s41077-016-0025-y>.

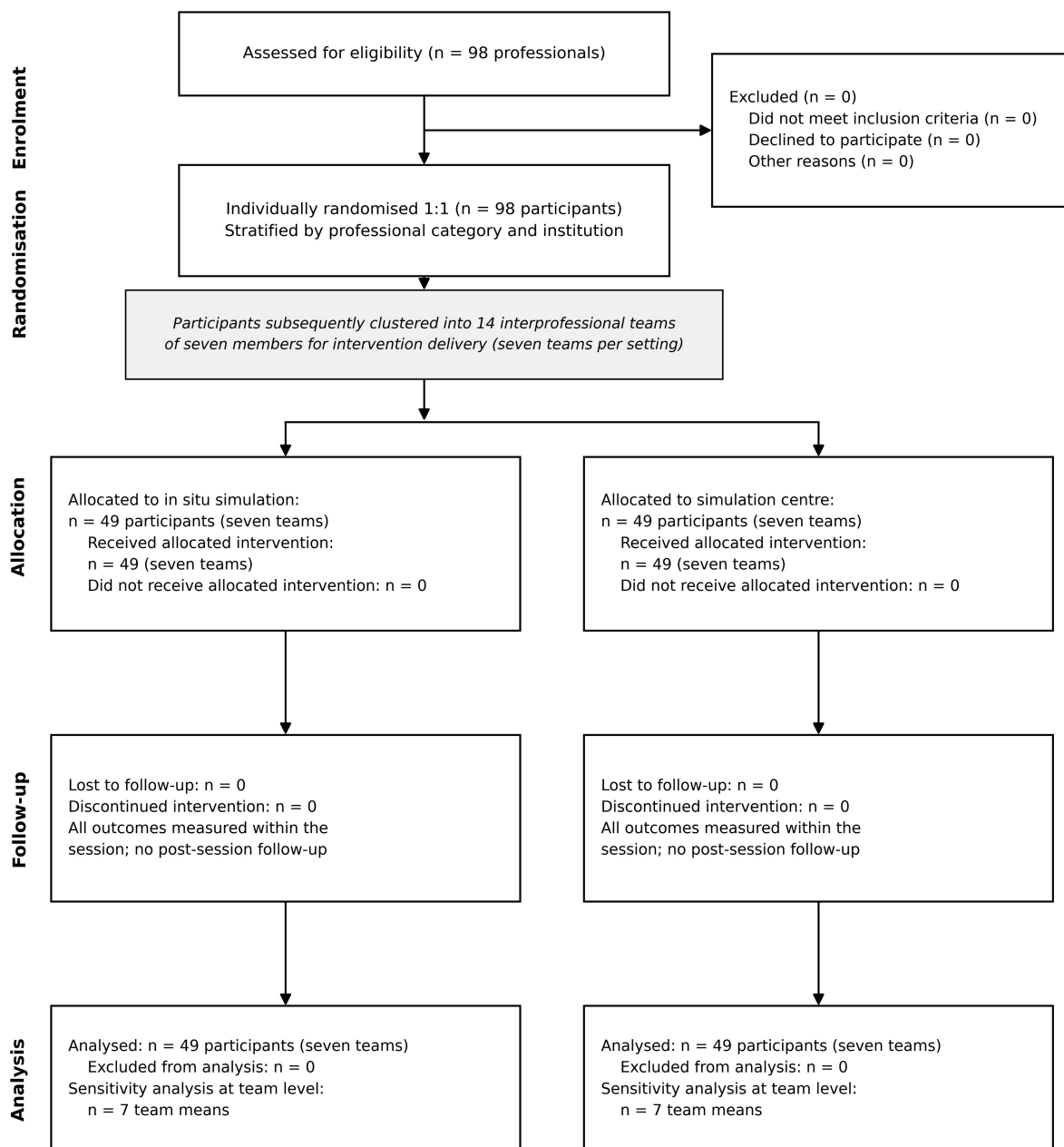
40. Pisciotanni F, Ramos-Magalhães C, Figueiredo AE. Efeitos da aplicação periódica da simulação in situ para educação permanente em ressuscitação cardiopulmonar no contexto da hemodiálise. *Enferm Nefrol.* 2020;23:274–84. <http://doi.org/10.37551/s2254-28842020029>.
41. Boet S, Bould MD, Fung L, Qosa H, Perrier L, Tavares W, et al. Transfer of learning and patient outcome in simulated crisis resource management: a systematic review. *Can J Anaesth.* 2014;61(6):571–82. <http://doi.org/10.1007/s12630-014-0143-8>.
42. Edmondson A. Psychological safety and learning behavior in work teams. *Adm Sci Q.* 1999;44(2):350–83. <http://doi.org/10.2307/2666999>.
43. Rudolph JW, Raemer DB, Simon R. Establishing a safe container for learning in simulation: the role of the presimulation briefing. *Simul Healthc.* 2014;9(6):339–49. <http://doi.org/10.1097/SIH.0000000000000047>.
44. Kolbe M, Eppich W, Rudolph J, Meguerdichian M, Catena H, Cripps A, et al. Managing psychological safety in debriefings: a dynamic balancing act. *BMJ Simul Technol Enhanc Learn.* 2020;6(3):164–71. <http://doi.org/10.1136/bmjstel-2019-000470>.

Figures



Figure 1

Visual and functional comparison of the two physical settings. The perioperative service where the in situ sessions took place and the simulation centre (SIMFMUP) room configured beforehand to replicate its layout and equipment.



Randomisation was at the level of the individual professional. Teams were constituted after allocation and the intervention was delivered to the team, so participants are clustered within teams inside each study setting (individually randomised group-treatment design). The unit of analysis was the individual participant.

Figure 2

Flow of participants through the trial (CONSORT 2025). Professionals were randomised individually and grouped after allocation into teams of seven for delivery of the intervention; all analyses were at participant level.

Supplementary Files

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

- [AdditionalFile6Changestothe trialprotocol.docx](#)
- [AdditionalFile4Structuredfacilitatorsobservationgrid.docx](#)
- [AdditionalFile5LASTknowledgequestionnaire.docx](#)
- [AdditionalFile3StandardisedPEARLSdebriefingscript.docx](#)
- [AdditionalFile1LASTclinicalscenario.docx](#)
- [Additionalfile8Reportingofthesimulationbasedintervention.docx](#)
- [AdditionalFile2LASTsimulation sessionplan.docx](#)
- [AdditionalFile7CompletedCONSORT2025reportingchecklist.docx](#)