Weagle et al. BMC Medical Education (2026) 26:868
https://doi.org/10.1186/s12909-026-09189-5

BMC Medical Education

RESEARCH

Open Access

Immersive learning in quality improvement: evaluating a virtual reality escape room for medical learners and faculty

Rachael Weagle[1], Maddie J. Venables[1] and Kheira Jolin-Dahel[1,2,3*]

Abstract

Background Didactic approaches to quality improvement (QI) are often perceived as insufficiently engaging, which may limit learner motivation and application of QI concepts. We developed a virtual reality (VR) escape room to reinforce core QI principles and conducted a Phase 1 evaluation focused on feasibility, usability, and perceived educational utility among medical learners and faculty.

Methods We designed a 60-90-minute multi-puzzle VR escape room aligning with QI principles (terminology, SMART aims, root cause analysis, change ideas, Plan-Do-Study-Act cycles, run charts). Twenty-seven participants (undergraduate medical students, residents, faculty) formed teams of three to four. We collected post-session surveys (Likert, yes/no, free-text) and concurrent "think-aloud" observations during gameplay to characterize usability and perceived educational utility. Survey data were collected using a standardized questionnaire administered either electronically or verbally with items read verbatim and responses recorded without prompting. Free-test survey responses and think-aloud field notes were analyzed using a consistent inductive, theme-based approach, and closed-ended survey items were summarized descriptively. This evaluation was intended to inform iterative refinement of the intervention.

Results Overall enjoyment was high (mean 3.9/5). Twenty-three of 27 participants (85%) reported that the VR escape room was perceived as helpful in reinforcing previously taught QI concepts. Analysis of survey free-text responses identified three themes: strengths of the experience (teamwork, engagement, collaborative problem solving), usability challenges (motion sickness, session length), and design preferences (variation in puzzle clarity and perceived linkage to QI concepts). Think-aloud findings similarly highlighted immersion, discussion, and collaboration as positive features, while hardware constraints and physical discomfort limited usability for some participants. Participant feedback informed iterative refinements related to puzzle design, facilitation, and technical setup.

Conclusion This Phase 1 pilot demonstrated that a VR escape room is a feasible and acceptable adjunct to existing QI teaching, providing learners with an immersive opportunity to apply previously introduced concepts. The main contribution of this study is feasibility and usability evidence, as well as implementation considerations (space, staffing, tolerability, and technical setup) to support educators considering immersive QI reinforcement activities. As

*Correspondence:
Kheira Jolin-Dahel
kdahe075@uottawa.ca
Full list of author information is available at the end of the article

objective knowledge outcomes were not assessed, future work will evaluate educational effectiveness using pre/post testing and comparative non-VR conditions following further refinement and curricular integration.

Keywords Quality improvement, Medical education, Virtual reality, Escape room, Gamification, Feasibility study

Backgrounds

Immersive group learning through virtual escape rooms offers a novel approach to enhance medical education. Traditional didactic learning methods, such as PowerPoint lectures, often fall short in consistently engaging learners and fostering knowledge retention [1]. As such, we designed a virtual reality (VR) escape room for medical learners and faculty, specifically targeting consolidation of quality improvement (QI) knowledge. This innovation was piloted to assess its feasibility, usability, and perceived educational utility as an adjunct to existing QI teaching.

Over the past decade, VR has become an increasingly popular educational tool, with predictions that it may evolve into a dominant computing platform [2]. Past reviews suggest that VR increases learner engagement and motivation, thereby improving retention of information [3]. This is thought to result from the positive emotions associated with immersive learning experiences [4]. VR is adaptable to a range of educational contexts across undergraduate, postgraduate, and health professions education. Recent reviews indicate that VR in medical education is evolving beyond procedural training, with growing evidence supporting its role in developing clinical decision-making, communication, and other non-technical skills [5, 6]. In these settings, it may provide an engaging complement to traditional approaches for teaching QI principles. In parallel, an expanding body of literature demonstrates the educational potential of escape rooms, both physical and virtual, highlighting their ability to promote teamwork, engagement, and, in some cases, knowledge gains [7–11]. However, reviewers have noted that outcomes vary with puzzle design quality, clarity of learning objectives, and alignment between puzzles and instructional goals. This suggests that immersive format alone is insufficient, and that the educational utility of such approaches depends on careful instructional design and alignment with intended learning objectives.

In this context, QI represents a key area in which such educational approaches may be particularly relevant across the continuum of medical education. Accreditation and competency frameworks increasingly emphasize physician participation in QI and patient safety activities across training and practice [12–15]. Learners and practicing clinicians are increasingly expected to understand core QI concepts and participate in improvement work. Despite this expectation, QI remains challenging to teach. Evidence from systematic reviews of QI education suggests that experiential, applied, and contextually grounded approaches, particularly those involving real-world QI projects, are associated with improved learner outcomes, whereas didactic teaching alone is insufficient to achieve higher-level outcomes [16–18]. Higher-level outcomes include changes in learner behaviour and the application of skills in practice, as well as downstream impacts on organizational or system-level performance outcomes, consistent with established evaluation frameworks [19]. Many learners find QI concepts methodologically unfamiliar. QI is also often introduced through didactic formats that may not adequately support application, discussion, and retention. Educational strategies that promote active, collaborative use of QI concepts may therefore be useful as complements to foundational teaching.

At the University of Ottawa, exposure to QI principles occurs across learner groups through introductory didactic teaching and participation in QI-related activities, although the depth and structure of that exposure vary by training level and role. Building on this need, our team developed a VR escape room for medical learners and faculty as a reinforcement activity rather than a primary instructional intervention. The escape room requires participants, in teams of three to four, to solve ten puzzles designed to reinforce QI concepts such as terminology, SMART aims, root cause analysis, change ideas, Plan-Do-Study-Act (PDSA) cycles, and run charts. The tool is intended for learners with some prior exposure to QI and is designed to support consolidation and application of concepts in a collaborative setting. Accordingly, as an early-stage pilot, this study focused on feasibility, usability, and perceived educational utility.

Methods

Study design and aim

We evaluated the learner experience of a VR escape room developed as a complementary teaching strategy to reinforce QI principles among medical learners and faculty at the University of Ottawa. Using a mixed-methods design, we collected post-session surveys and concurrent think-aloud observations during gameplay to assess feasibility, usability, and perceived educational utility, as well as identify implementation considerations.

As an early-stage pilot, the primary aim was to determine whether the intervention could be implemented as intended, was acceptable to learners, and was perceived as educationally useful as an adjunct to existing QI teaching. Objective learning outcomes were not assessed in this phase. This study was a Phase 1 feasibility and usability evaluation intended to inform iterative refinement of the intervention prior to future evaluation of learning outcomes.

Intervention: VR escape room and alignment to QI concepts

The intervention included a one-hour introductory didactic session on QI principles delivered prior to the VR escape room. This session introduced foundational QI concepts that were subsequently reinforced within the escape room activities. The VR escape room was designed for learners with beginner-level QI knowledge and positioned as an application and reinforcement activity rather than as primary instruction.

The VR escape room provided participants with a 60 to 90-minute immersive experience in which teams of three to four completed ten puzzles designed to apply knowledge from their prior QI training. The puzzles addressed core QI domains including terminology, SMART aims, root cause analysis methods such as process mapping, fishbone diagrams, and Pareto charts, change ideas, PDSA cycles, and run charts. Participants navigated multiple virtual rooms with 3D visualization and sound effects, working collaboratively to solve each puzzle. Each puzzle was intentionally mapped to a specific QI learning objective to support constructive alignment between gameplay mechanics and educational content. Table 1 outlines the puzzles and their alignment with QI principles.

Representative puzzle activities illustrated application of core QI tools within the immersive environment. For example, in the process mapping activity, participants assembled standardized process map symbols to reconstruct a clinical workflow sequence. In the Pareto chart exercise, teams reorganized frequency data to identify priority drivers of a clinical problem. Participants also prioritized proposed change ideas based on Pareto findings and activated a virtual PDSA button, reinforcing the progression from problem analysis to intervention testing. Finally, run chart interpretation tasks required learners to assess time-series data to determine the impact of implemented change. Collectively, these activities reinforced applied use of foundational QI methods introduced during the didactic QI session.

Practical implementation and facilitation

Sessions were delivered in a dedicated simulation room with adequate space for safe movement. One facilitator oriented participants, monitored safety, and managed session flow, while a second staff member assisted with headset fitting, troubleshooting, and sanitation. Headsets were pre-charged, and spare controller batteries and lens spacers were available for participants wearing glasses. A brief pre-brief introduced QI objectives and headset controls, followed by a structured debrief reinforcing key concepts. A designated observer documented think-aloud field notes and did not participate in facilitation or data collection beyond observation.

Participants

Participants included undergraduate medical students, family medicine residents, and faculty affiliated with the Department of Family Medicine at the University of Ottawa. A total of 27 participants took part in the study (19 undergraduate medical students, 4 family medicine residents, 4 faculty).

We recruited participants through email invitations distributed within the department. Invitations were sent to approximately 328 undergraduate medical students and 170 residents, as well as faculty through departmental networks. Participation was voluntary and based on convenience sampling, with enrollment occurring on a first-come, first-served basis until session capacity was reached. Seven sessions were delivered, each accommodating three to four participants. No incentives were offered.

All participants received introductory QI teaching as part of the intervention; prior exposure beyond this session varied and was not formally measured, consistent with the feasibility focus. Participants were assigned to teams of three to four, with mixed training levels when feasible to promote collaborative problem solving. Team composition was not fully standardized due to scheduling constraints.

Data collection

We collected data from two sources: (1) post-session surveys and (2) think-aloud observational field notes recorded during gameplay.

Post-session surveys

Immediately following the VR escape room, participants completed an individual survey capturing their perceptions of the experience. The survey was co-developed with faculty and learners and piloted for clarity (Supplementary File). It included (i) Likert-scale items (0 to 5) assessing overall enjoyment and perceived difficulty; (ii) yes/no questions on enjoyment and perceived knowledge consolidation; and (iii) three open-ended questions on usability, educational value, and puzzle design.

We administered surveys electronically or verbally using the same standardized survey instrument, with questions read verbatim and responses recorded without prompting to ensure consistency across participants. Only responses obtained through this standardized instrument were included in the analysis. No additional

Table 1 Details on the QI virtual reality experience for medical learners and faculty

Puzzle	QI principle being addressed	Applied learning objectives
Puzzle 1.1: Build the QI map	Brief review of steps of a QI project	Participants search the virtual environment to locate scattered jigsaw puzzle pieces. Working collaboratively, they assemble the pieces to reveal a completed image corresponding to a slide previously presented during the introductory QI lecture outlining the sequential steps of a QI project (e.g., identifying a problem, defining an aim, selecting measures, testing change, and evaluating results). This opening puzzle orients learners to the VR controls while reinforcing recognition of the overall structure and flow of a QI initiative.
Puzzle 1.2: Find the gap	Gap analysis	Participants locate three short written clinical scenarios within the virtual environment. Each scenario represents a different concept: a true gap in care, a guideline recommendation, or a general consensus statement. Working as a team, participants identify which scenario reflects a measurable gap between current and desired performance. They then match the selected paragraph to a corresponding "answer key" displayed on the wall. Correct placement highlights specific letters within the text, which together form a code used to unlock a safe. This puzzle reinforces the distinction between identifying a performance gap and referencing guidelines or general best practices.
Puzzle 1.3: Scrabble	Review of QI terminology	Participants interact with a Scrabble-style table displaying scrambled letter tiles. For each term, the team is provided with a brief definition of a QI concept and must rearrange the mixed letters to spell the correct term (e.g., benchmark, team, measure, quadruple aim). Successful completion of each word unlocks progression to the next step in the puzzle. This activity reinforces recall and accurate use of foundational QI terminology introduced during prior instruction.
Puzzle 1.4: Monkeys	SMART AIM	Participants locate five "monkey" statues, each representing one component of a SMART aim. They are also presented with five project statements, each intentionally missing one SMART element. The puzzle is positioned at the top of a virtual ladder, requiring coordination between team members positioned at different levels of the environment. One participant identifies and passes components from below while another matches the SMART element to the appropriate incomplete statement above. The puzzle unlocks once all matches are correct. This design promotes communication and collaborative problem-solving while reinforcing recognition of the distinct components required to construct a well-defined SMART aim.
Puzzle 2.1: Build your team	Root cause analysis tool: process map	Participants are presented with a collection of numbered, building block style pieces placed on a virtual table. Each block corresponds to a standard process map symbol (e.g., circle for start or end, rectangle for process step, or diamond for decision point). Multiple blocks share the same number while differing in shape. On the wall, participants see a written outline describing the sequence of a patient visit in clinic, including variations when a learner is present versus absent. The team selects the correctly shaped block for each numbered step and places it on a blank process map template. The puzzle is solved when the sequence and symbols are correctly assembled, reinforcing correct use of process map conventions and structured visualization of clinical workflows.
Puzzle 2.2: The safe	Root cause analysis tool: worst possible idea	Participants encounter a virtual wall displaying a collection of exaggerated or intentionally poor "change ideas" related to a clinical problem. Embedded within the image are hidden numbers that form the code to unlock a nearby safe. To identify the correct code, participants must recognize that the displayed ideas represent the "Worst Possible Idea" brainstorming technique, a structured method used in root cause analysis to surface assumptions and common pitfalls. Once the team extracts the hidden numbers, they enter the code to advance. This puzzle reinforces familiarity with creative brainstorming approaches used during early-stage QI planning.
Puzzle 2.3: Tilt table	Root cause analysis tool: Fishbone diagram	Participants encounter an interactive "tilt table" designed to resemble a fishbone diagram. They collect a set of balls, each labeled with a potential root cause related to a clinical scenario. Working in pairs, participants tilt and balance the table to guide each ball into the appropriate category (e.g., people, process, environment, equipment). Completion occurs when all root causes are accurately categorized, reinforcing structured cause analysis within the fishbone framework.
Puzzle 3.1: The QI lab	Root cause analysis tool: Pareto chart	Participants enter a virtual "QI lab" where they encounter a digital display showing a bar graph of problem frequencies arranged in no particular order. The team recognizes that the graph represents a Pareto chart and reorganizes the bars in descending order of frequency to identify the most significant contributing factors. Once the data are correctly ordered, the prioritized issue becomes visually apparent and unlocks the next step. This activity reinforces recognition and interpretation of Pareto charts to support data-informed prioritization in QI initiatives.

Table 1 (continued)

Puzzle	QI principle being addressed	Applied learning objectives
Puzzle 3.2: Make a change	Prioritizing change ideas	This puzzle builds directly on the preceding Pareto chart activity. Participants locate virtual clipboards labeled with a proposed change idea corresponding to one of the root causes identified previously. Using the prioritized order established from the Pareto chart, the team arranges the change ideas in descending order of impact. The Pareto display and change idea clipboards are located in separate areas of the virtual environment, requiring participants to retrieve information or collaborate across spaces. To complete the puzzle, participants correctly order the change ideas and activate a virtual "PDSA" button symbolizing the transition from analysis to testing change. The puzzle unlocks once both the prioritization and PDSA activation steps are completed, reinforcing the linkage between data-driven prioritization and iterative improvement cycles. This activity emphasizes the translational step from problem identification to the selection of an intervention and highlights the importance of prioritizing feasible, high-impact changes when designing QI initiatives.
Puzzle 3.3: Run with it	Run chart	Following activation of the virtual PDSA cycle, a run chart appears displaying time-series data reflecting the impact of the implemented change. Participants interpret the direction and pattern of the data over time (e.g., improvement, decline, or stability). The direction of change corresponds to the sequence required to unlock a directional lock mechanism. The puzzle is completed when the team correctly interprets the run chart and enters the appropriate sequence, reinforcing recognition of temporal trends and the use of run charts to assess change over time.

All puzzles were designed to reinforce concepts introduced during the introductory QI didactic session and to support collaborative application of foundational improvement tools

verbal or conversational data were collected or analyzed as a separate data source.

Brief clarification conversations occasionally occurred following survey completion to ensure accurate interpretation of responses; however, these interactions were informal, not systematically recorded, and were not included in the analysis.

Think-aloud observations

During gameplay, participants were encouraged to verbalize their thoughts, reactions, and decision-making processes as they interacted with the VR environment. This think-aloud method is widely used in usability testing and educational tool evaluation [20, 21].

Observations were captured as handwritten field notes by a designated observer. The observer documented participant utterances and behaviours related to usability, navigation, collaboration, hardware interaction, and puzzle engagement without providing instructional guidance. Field notes focused on contemporaneous verbalizations during gameplay and were used to contextualize survey findings.

Data analysis

We analyzed qualitative data from survey free-text responses and think-aloud field notes together using a consistent inductive, theme-based approach in NVivo. This enabled identification of recurring themes and patterns in participant experiences, as well as suggestions for tool refinement.

We summarized survey responses using descriptive statistics. No inferential testing was conducted given the exploratory scope and sample size. One author (RW) conducted initial coding, and a second author (KJD) reviewing coded excerpts and thematic interpretations to assess their alignment with the source data. Discrepancies were resolved through discussion and iterative refinement, supporting analytic rigour and confirmability.

Ethics

As a QI study, this work was deemed exempt by the Health Sciences and Science Research Ethics Board at the University of Ottawa, permitting verbal consent procedure. Participants provided verbal consent for participation and anonymous use of their feedback and were informed that participation was voluntary and that they could withdraw at any time.

Results

We recruited a total of 27 participants, including 19 undergraduate medical students, four family medicine residents, and four faculty, between July and October 2022. The participants were grouped into seven teams of three to four individuals. The post-session survey response rate was 100% among those who participated (27/27). Given the exploratory scope and small sample size of this pilot, we limited the analyses to descriptive statistics.

Survey outcomes

Of the 27 participants, none (0%) rated their enjoyment of the VR escape room as low. Four participants (15%) rated it as "okay", four (15%) were neutral, ten (37%) reported that they enjoyed the experience, and nine (33%) described it as "amazing". The mean enjoyment score was 3.9 out of 5. One participant reported that the experience was "much too easy", while another rated it as "slightly easy". All remaining participants reported that the level of difficulty was appropriate.

When asked whether the VR escape room was perceived as helpful in consolidating previously introduced QI knowledge, 23 participants (85%) responded "yes".

Qualitative analysis of survey free-text responses

Qualitative analysis of free-text survey responses identified three key themes: (1) strengths of the experience, (2) usability challenges, and (3) design preferences.

Strengths of the experience

Participants frequently highlighted teamwork, engagement, and opportunities for collaborative problem-solving as key strengths of the VR escape room. Many participants described the experience as interactive and conducive to applying previously introduced QI concepts. Participants also noted that collaboration facilitated discussion and reinforcement of QI concepts during gameplay.

Usability challenges

Several participants (15%) reported nausea, attributing symptoms to the duration of gameplay or abrupt movements in the VR environment. One participant commented, "Shorter session; it was too long and made me nauseous" (Participant #4). Another noted, "Sharp horizontal changes were very nauseating; I would prefer smoother transitions" (Participant #7). Participants suggested that smoother visual transitions and shorter periods of continuous headset use could improve tolerability.

Design preferences

Participants identified variability in puzzle clarity and alignment with QI concepts. Some participants expressed a desire for stronger conceptual linkage between puzzles and explicit reinforcement of QI principles. For example, one participant stated, "For the first puzzle in the last room, it would be nice if the connection to the previous puzzle was made more obvious. Some puzzles did not have to do with QI and took too long." (Participant #15).

Preferences for specific puzzles were also reported. The Pareto chart puzzle (22%) and gap analysis puzzle (15%) were most frequently identified as least preferred, while the process mapping puzzle (52%) and Scrabble puzzle (37%) were most frequently cited as preferred activities. These findings informed iterative refinement of puzzle clarity, instructional scaffolding, and alignment with intended learning objectives. The findings from survey responses were consistent with themes identified in think-aloud observations.

Think-aloud outcomes

Think-aloud observations captured participants' contemporaneous verbalizations during gameplay, focusing on usability, collaboration, and interaction with the VR environment and puzzles. The analysis of field notes yielded three overarching themes: learner experience, hardware considerations, and administration and player dynamics.

Learner experience

Most participants (85%) described their overall experience as positive, citing teamwork, social bonding, and opportunities to apply previously introduced QI concepts. Four participants (15%) reported a negative experience, primarily due to nausea or physical discomfort during gameplay. Negative experiences were not attributed to lack of prior QI knowledge. One participant commented, "Yes very much - it allowed me to use the QI concepts, which really helped me consolidate what we learned in the lecture." (Participant #19). Another shared, "Yes, especially when we could work with each other to remind ourselves of the QI strategy." (Participant #22).

Hardware

The feedback related to hardware was consistently critical and formed a distinct usability theme. The reported challenges included difficulty wearing glasses within the headset, controller battery depletion, limited charging access, and system timeouts if gameplay was paused. One participant commented, "The hand controllers eat through batteries very quickly." (Participant #15).

Administration and player dynamics

The team-based structure facilitated interaction and active discussion. Noise levels and physical space constraints were frequently noted. While most groups described the noise as distracting, one group perceived it as beneficial to collaboration, stating, "It got loud at times, but that helped with team bonding." (Participant #6).

Discussion

Integration of findings

Taken together, survey and think-aloud data suggested that participants perceived the VR escape room as engaging and acceptable approach to applying and reinforcing previously introduced QI concepts. Positive experiences were associated with teamwork and perceived appropriateness of difficulty, while negative experiences were linked to motion sickness, session length, and technical limitations. These findings informed iterative refinements to the tool, including smoother visual transitions, shorter continuous headset intervals, and strengthened alignment between puzzles and QI learning objectives. The findings from this Phase 1 evaluation are therefore interpreted as indicators of feasibility, usability, and participant-reported educational utility.

We designed the VR escape room as a complementary educational strategy to reinforce QI principles among medical learners and faculty. Consistent with its Phase 1 feasibility focus, the study examined usability, acceptability, and perceived educational utility without assessing objective learning outcomes. The primary purpose was to identify strengths, limitations, and implementation considerations to inform iterative refinement of the escape room prior to curricular integration and future evaluation of learning outcomes.

The analysis revealed variability in overall learner experience. Positive outcomes were associated with teamwork, collaborative problem solving, and appropriate difficulty, while negative outcomes were linked to motion sickness, session length, and technical limitations. These findings are consistent with prior reviews of VR, including medical education, which have highlighted challenges such as technical limitations, difficulties simulating real-world complexity, session-related length issues, and cybersickness [5, 22]. Notably, effective team communication and opportunities for collaborative discussion appeared to contribute to more positive learner experiences. Given that all participants completed an introductory QI session prior to gameplay, the escape room functioned as an application environment in which previously introduced concepts could be revisited. Participants who reported familiarity with the introductory QI concepts appeared to navigate the VR experience with greater ease, suggesting that positioning the escape room as a reinforcement activity following foundational instruction may optimize learner experience.

These observations suggest that pairing the VR escape room with a proximal didactic QI lecture and structured facilitation may optimize its perceived educational value, consistent with prior studies emphasizing the importance of facilitation, guidance, pre-briefing, and complementary teaching [7, 23]. However, these interpretations should be understood as hypothesis-generating and require validation through studies assessing objective outcomes.

Conversely, negative experiences were primarily linked to motion sickness and nausea, particularly among participants predisposed to motion-related malaise. Similar reports of cybersickness have been described in the broader VR medical education literature [5, 24–26]. Potential solutions include shortening game duration, encouraging player movement during gameplay, and eliminating nausea-inducing control features [26]. Future iterations of VR educational tools should incorporate explicit cybersickness mitigation strategies to improve comfort and learner engagement.

Additionally, participant feedback highlighted puzzles requiring refinement, including the Pareto chart, SMART aim, and Fishbone diagram. Planned improvements include providing clarifying instructions, integrating hint systems, and strengthening links between puzzles and QI concepts. In contrast, the process mapping and Scrabble puzzles were consistently rated as favourites, reflecting the importance of intuitive design and clear alignment with QI principles. In their dual-site case study, Beger et al. [27] similarly reported that puzzles and riddles were the most engaging aspects of VR escape rooms, as they encouraged students to "think about material in a new way" and identify knowledge gaps, whereas poorly designed tasks hindered engagement. These findings underscore the importance of constructive alignment between puzzle mechanics, cognitive load, and intended QI concepts, particularly when designing immersive tools intended to support application and reinforcement rather than primary instruction [27]. Collectively, these observations highlight the need for strong alignment between puzzle mechanics, learner cognitive load, and intended QI learning objectives.

Strengths and limitations

This pilot offers several strengths, including the use of both surveys and think-aloud observations to capture complementary perspectives on learner experience, and the inclusion of undergraduate medical students, residents, and faculty, which provided a broad view of usability across learner levels. The VR escape room was implemented in a realistic teaching environment, supporting ecological validity and generating practical insights applicable to program delivery. However, the study also has limitations. The small convenience sample and single-institution context limit generalizability, and the absence of objective pre/post knowledge testing restricts conclusions regarding learning outcomes or impact on practice. Learners differed in their prior exposure to QI concepts, contributing to variability in engagement and performance. These limitations are consistent with the early-stage, feasibility-focused scope of the project and informed the design of the next phase of evaluation.

In addition, qualitative data were derived from both survey free-text responses and think-aloud observations, which, while analyzed using a consistent inductive approach, may reflect context-specific participant experiences and are not intended to represent transferable outcomes without further validation.

While this study demonstrates feasibility and perceived educational utility, the resource-intensive nature of VR implementation, including hardware requirements, physical space, and facilitation, may limit reproducibility across settings. Although the VR escape room is openly available for use and adaptation, variability in local curricular structures, technical infrastructure, and access to immersive technology may affect generalizability.

As such, this work is best understood as an innovation-focused feasibility study intended to inform adaptation and future evaluation of outcomes rather than a directly generalizable educational intervention.

Overall, this pilot demonstrated that a VR escape room can engage medical learners and faculty in consolidating QI knowledge. These findings should be interpreted as evidence of feasibility and perceived utility rather than evidence of effectiveness. While the project is in its early stages, these findings provide important guidance for refinement and support the educational value of the tool. Given that this phase focused on feasibility and perceived learning, objective knowledge outcomes were not collected, and future phases will address this limitation. The next steps will focus on iterative improvements, integration into the family medicine residency curriculum, and broader implementation. Fusco et al. [28] demonstrated that integrating escape rooms with simulation improved readiness for teamwork and suggested value for curriculum integration. Similarly, Rafi et al. [29] highlight the potential for VR and gamified escape rooms across multiple specialties and learner levels, while emphasizing that more rigorous studies are needed to confirm effectiveness.

The VR escape room is openly available for use and adaptation online through eCampus Ontario: https://openlibrary-repo.ecampusontario.ca/jspui/handle/123456789/1094. With iterative refinement, wide-scale deployment has the potential to enhance QI education by offering an innovative and engaging complement to traditional approaches, while recognizing that not all learners may prefer this format.

Conclusion

This exploratory pilot demonstrated that a VR escape room is a feasible and engaging adjunct for consolidating QI principles among medical learners and faculty. Participants valued the teamwork and immersive problem solving inherent in the format, and qualitative feedback identified key design and usability features that influence learner experience. However, this study did not assess objective learning outcomes, and therefore no conclusions can be drawn regarding impact on knowledge, behaviour, or practice. Several usability issues such as cybersickness, session duration, and puzzle clarity require refinement prior to broader implementation.

From a feasibility perspective, this study also identified several practical considerations relevant to implementation and scalability. These included the need for dedicated physical space, multiple facilitators to support technical troubleshooting and participant safety, advance preparation of hardware, including headset charging and spare batteries, and attention to learner comfort and tolerability. While hard costs were not formally analyzed, faculty time for development, facilitation, and iteration represents a nontrivial investment that must be weighed when considering broader adoption. These considerations are critical for educators evaluating whether immersive technologies are feasible within their local contexts.

Overall, this Phase 1 evaluation demonstrated that a VR escape room can be implemented as a feasible and engaging adjunct to existing QI teaching, providing learners with an opportunity to collaboratively apply previously introduced concepts in an immersive environment. These findings primarily inform implementation and design considerations rather than educational outcomes. Future work will build on this feasibility foundation by evaluating learning outcomes and comparative effectiveness within a structured curriculum. Such studies will be required to determine the extent to which this intervention influences knowledge, behaviour, or practice.

Abbreviations

QI	Quality improvement
VR	Virtual reality
PDSA	Plan-Do-Study-Act

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12909-026-09189-5.

Supplementary Material 1.

Acknowledgements

We thank the Department of Family Medicine staff who supported logistics and headset preparation, and the learners and faculty who participated in the QI pilot study.

Use of large language models

We used a large language model-based assistant (ChatGPT; OpenAI) for language editing and formatting. No content generation replaced author analysis or interpretation, and all authors reviewed and approved the final text.

Authors' contributions

RW conceived the study, coordinated implementation, and contributed to analysis and interpretation. MJV contributed to study design, data analysis, and manuscript drafting. KJD supervised the project, contributed to conception and design, interpretation, and critical revision. All authors read and approved the final manuscript.

Funding

This project was funded by eCampus Ontario, Digital Content, Create a New Simulation, Serious Game or XR Experience Grant (OTTA-1262).

Data availability

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This project was conducted in accordance with international guidelines and regulations including those in the Declaration of Helsinki and the ones followed at the University of Ottawa. As a QI study, this work was deemed exempt by the Health Sciences and Science Research Ethics Board at the University of Ottawa on December 8, 2021. The exemption determination permitted verbal consent procedures for participation and dissemination of anonymized findings. Participants were invited to complete a VR escape room experience, a post-session survey, and a think-aloud session. Participants were informed that participation was voluntary, that they could withdraw at any time, and that de-identified feedback from surveys and think-aloud observations could be used for publication. Verbal consent was obtained prior to participation by the senior author (KJD).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Family Medicine, University of Ottawa, 201-600 Peter Morand Crescent, Ottawa, ON K1G 5Z3, Canada
[2]Director of Quality Improvement and Family Medicine Resident Scholarly Project Programs, Ottawa, Canada
[3]Family Physician, Nation River Clinic, Winchester, ON, Canada

Received: 7 November 2025 / Accepted: 7 April 2026
Published online: 18 April 2026

References

1. Campos E, Hidrogo I, Zavala G. Impact of virtual reality use on the teaching and learning of vectors. Front Educ. 2022;7:965640. https://doi.org/10.3389/feduc.2022.965640.
2. Makransky G, Lilleholt L. A structural equation modeling investigation of the emotional value of immersive virtual reality in education. Educ Tech Res Dev. 2018;66(5):1141–64.
3. Sviridova E, Yastrebova E, Bakirova G, Rebrina F. Immersive technologies as an innovative tool to increase academic success and motivation in higher education. Front Educ. 2023;8:1192760. https://doi.org/10.3389/feduc.2023.1192760.
4. Holzinger A. Usability engineering methods for software developers. Commun ACM. 2005;48(1):71–4.
5. Mergen M, Graf N, Meyerheim M. Reviewing the current state of virtual reality integration in medical education-a scoping review. BMC Med Educ. 2024;24(1):788.
6. Fernández-Alcántara M, Escribano S, Juliá-Sanchis R, Castillo-López A, Pérez-Manzano A, Macur M, et al. Virtual simulation tools for communication skills training in health care professionals: literature review. JMIR Med Educ. 2025;11(1):e63082.
7. Quek LH, Tan AJ, Sim MJ, Ignacio J, Harder N, Lamb A, et al. Educational escape rooms for healthcare students: a systematic review. Nurse Educ Today. 2024;132:106004.
8. He Y, Chen X, Ma H, Zhao R, Zhou H, Yang Y. Effectiveness of escape room in Medical Education: a systematic review and meta-analysis. Games Health J. 2024;13(6):419–27.
9. Moore L, Campbell N. Effectiveness of an escape room for undergraduate interprofessional learning: a mixed methods single group pre-post evaluation. BMC Med Educ. 2021;21(1):220.
10. Hursman A, Richter LM, Frenzel J, Nice JV, Monson E. An online escape room used to support the growth of teamwork in health professions students. J Interprof Educ Pract. 2022;29:100545.
11. Kakos NJ, Lufler RS, Cyr B, Zwirner C, Hurley E, Heinrich C, et al. Unlocking knowledge: a meta-analysis assessing the efficacy of educational escape rooms in health sciences education. Adv Health Sci Educ. 2025;30(3):837–57.
12. Association of American Medical Colleges. Quality Improvement and Patient Safety Competencies Across the Learning Continuum. 2019. Available from: https://www.aamc.org/data-reports/quality-improvement-and-patient-safety-competencies-across-learning-continuum. Cited 24 Mar 2026.
13. Accreditation Council for Graduate Medical Education. Institutional Requirements. 2025. Available from: https://www.acgme.org/globalassets/pfassets/programrequirements/2025-reformatted-requirements/institutionalrequirements_2025_reformatted.pdf. Cited 24 Mar 2026.
14. Goldstein S, des Bordes JK, Neher S, Rianon N. Improving patient safety through mandatory quality improvement (QI) education in a family medicine residency programme. BMJ Open Qual. 2024. https://doi.org/10.1136/bmjoq-2023-002623.
15. Hamline MY, Rasmussen LN, Ratanasen M, Li STT. Teaching quality improvement by doing quality improvement: an experiential medical education curriculum to improve didactic attendance. Acad Pediatr. 2026. https://doi.org/10.1016/j.acap.2026.103248.
16. Peiris-John R, Selak V, Robb G, Kool B, Wells S, Sadler L, et al. The state of quality improvement teaching in medical schools: a systematic review. J Surg Educ. 2020;77(4):889–904.
17. Lavin J, Kao LS. Building infrastructure to teach quality improvement. Semin Pediatr Surg. 2023;32(2):151283. https://doi.org/10.1016/j.sempedsurg.2023.151283.
18. Herman DD, Weiss CH, Thomson CC. Educational strategies for training in quality improvement and implementation medicine. ATS scholar. 2020;1(1):20–32.
19. Kirkpatrick DL. Seven keys to unlock the four levels of evaluation. Perform Improv. 2006;45(7):5–8.
20. Barnum CM. Usability testing essentials: Ready, set... test! Morgan Kaufmann; 2020.
21. Fonteyn ME, Kuipers B, Grobe SJ. A description of think aloud method and protocol analysis. Qual Health Res. 1993;3(4):430–41.
22. Sadek O, Baldwin F, Gray R, Khayyat N, Fotis T. Impact of virtual and augmented reality on quality of medical education during the COVID-19 pandemic: a systematic review. J Grad Med Educ. 2023;15(3):328–38.
23. Barrickman A, McMillan A, Gálvez-Peralta M, Purnell K. Development and assessment of integrated virtual escape rooms to reinforce cardiology content and skills. Am J Pharm Educ. 2023;87(3):ajpe8899.
24. Biswas N, Mukherjee A, Bhattacharya S. Are you feeling sick?–A systematic literature review of cybersickness in virtual reality. ACM-CSUR. 2024;56(11):1–38.
25. Simón-Vicente L, Rodríguez-Cano S, Delgado-Benito V, Ausín-Villaverde V, Delgado EC. Cybersickness. A systematic literature review of adverse effects related to virtual reality. Neurología (English Edition). 2024;39(8):701–9.
26. Kim HY, Kim EY. Effects of medical education program using virtual reality: a systematic review and meta-analysis. Int J Environ Res Public Health. 2023;20(5):3895.
27. Beger AW, Hannan S, Patel R, Sweeney EM. Virtual escape rooms in anatomy education: case studies from two institutions. Adv Physiol Educ. 2025;49(3):621–32.
28. Fusco NM, Foltz-Ramos K, Ohtake PJ. An interprofessional escape room experience to improve knowledge and collaboration among health professions students. Am J Pharm Educ. 2022;86(9):ajpe8823.
29. Rafi A, Habibi A, Kalantarion M. Traditional medical education under siege: the revolutionary integration of virtual reality and gamification in medical escape rooms. BMC Med Educ. 2025;25(1):898.

Publisher's note