

Enhancing home care safety: a randomized controlled trial of VR-based training for informal caregivers

Eva Gil-Hernández^{a,1} , Irene Carrillo^{b,1,*} , Clara Pérez-Esteve^a , Almudena Arroyo^c , Mercedes Guilabert^b , Purificación Ballester^d , José Joaquín Mira^{a,b} 

^a Grupo de Investigación ATENEA, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana (FISABIO), Alicante, Spain

^b Departamento de Psicología de Salud, Universidad Miguel Hernández de Elche, Elche, Spain

^c Fundación San Juan de Dios. Escuela Universitaria de Enfermería y Fisioterapia San Juan de Dios, Universidad Pontificia Comillas, Campus Bormujos (Sevilla), Spain

^d Facultad de Farmacia y Nutrición, UCAM, Universidad Católica de Murcia, Murcia, Spain

ARTICLE INFO

Keywords:

Virtual reality
Informal caregivers
Home care
Patient safety
Medication errors
Caregiver burden
Simulation-based training

ABSTRACT

Background: Europe's aging population and the move to home-based long-term care place growing demands on informal caregivers, who often lack formal training. This substantially increases both caregiving and medication errors and caregiver burden. Virtual reality (VR) enables experiential training but is seldom tailored to non-professional caregivers or evaluated in real-world conditions.

Objective: To analyze the ability of a brief training based on VR to reduce informal caregivers' burden and their caregiving/medication errors at home.

Methods: Two-arm randomized controlled trial in three Spanish regions. Informal caregivers were randomized to structured VR training or usual materials; N = 140 (70/70) caring for people with chronic conditions. Assessments at baseline and 3 months, aligned with the Kirkpatrick model: L1 satisfaction; L2 video-based error detection; L3 self-reported caregiving/medication errors; L4 emotional burden. The intervention delivered 18 immersive scenarios reflecting common home-care tasks.

Results: Satisfaction was high in the intervention arm ($\geq 90\%$ positive on usefulness, relevance, and applicability). Level 2: the intervention group improved error recognition in video scenarios (mean identified errors 5.41 to 6.64; mean change + 1.23; $P = 0.0001$), with 46/70 (65.7%) showing improvement ($\chi^2 = 33.114$; $p < 0.0001$). Level 3: self-reported errors decreased in the intervention group (62 to 23) but increased in controls (46 to 77); the time-by-group interaction was significant ($F = 11.53$; $P = 0.0009$). Level 4: emotional burden shifted toward lower categories at follow-up in the intervention group ($\chi^2 = 17.73$; $P = 0.0014$). Complementary measures showed an increase in COM-B total score from 6.38 to 7.43 ($P = 0.0017$), with improvements in Opportunity ($P = 0.0325$) and positive trends in Capability and Motivation.

Conclusions: A short, structured VR training improved recognition of unsafe practices and reduced self-reported caregiving/medication errors among informal caregivers, with concurrent reductions in emotional burden. Findings support integrating immersive, user-centered training into caregiver support programs to enhance the safety and quality of home care.

[ClinicalTrials.gov NCT05885334](https://clinicaltrials.gov/ct2/show/study/NCT05885334).

1. Introduction

1.1. Challenge

Europe is undergoing a profound demographic transformation. By 2050, the population aged 80 and over is expected to double [1],

increasing the demand for long-term care (LTC). At the same time, a growing majority of older adults express the desire to remain in their homes for as long as possible [2]. A key challenge is how to ensure that individuals receive care at home that meets quality standards comparable to those provided in formal health and social care institutions [3,4].

* Corresponding author at: Universidad Miguel Hernández de Elche, Avinguda de la Universitat d'Elx, s/n, Elche 03202, Spain.

E-mail address: icarrillo@umh.es (I. Carrillo).

¹ Joint first authors.

1.2. Caregiving errors

Informal caregivers commit errors while providing care at home [5–8]. Particularly as care tasks become more complex, which may increase the risk of adverse outcomes [9]. Such errors also contribute to heightened emotional distress and a greater sense of personal responsibility among caregivers, ultimately exacerbating their burden. There is an urgent need for accessible training programs aimed at enhancing the safety of home-based care [10].

1.3. Online support

To address this challenge, various digital approaches have been implemented to support informal caregivers. Online support groups, including moderated forums and mobile applications, offer psychosocial support, reduce isolation, and foster emotional validation through peer interaction [11–15].

1.4. Virtual reality

Virtual reality (VR) represents a promising yet underutilized tool to support informal caregivers. VR-based interventions have shown effectiveness in enhancing caregivers' practical skills, empathy, and emotional resilience [16–18]. VR enables rehearsal of safety-critical procedures—such as hygiene, mobilization, and medication management—within realistic and low-risk settings [13,16,19]. These improvements can contribute not only to patient safety but also to the caregivers' physical and psychological well-being and fostering a sense of competence [16,20].

Despite recent advances [21], the implementation of VR interventions continues to face significant challenges [12,22]. Few interventions simulate realistic home care scenarios or directly address frequent caregiver errors in routine tasks. These limitations underscore the need for more context-sensitive, rigorously evaluated VR training tools to improve caregiver competence and reduce preventable harm in home care [16,18,22,23].

2. Aim

This study aimed to analyze the ability of a VR-based training to reduce errors in the provision of care and medication at home by informal caregivers.

3. Methods

A two-arm experimental design compared an intervention group with a control group. Assessments were conducted at baseline, immediately post-intervention, and at three months. Participants were randomly assigned by simple randomization. The trial is part of RealityCare [24], which evaluates training and other disruptive technologies to improve safety for people with chronic conditions receiving home care from informal caregivers.

In this study, an informal caregiver was defined as any individual providing care at home to people with chronic illnesses, disabilities, or other long-term health or care needs, without holding formal qualifications or academic training in healthcare. Caregivers may be family members, close friends, or non-related individuals offering support within a personal relationship. This care may be either paid or unpaid [25]. An informal caregiver error was defined as any preventable event—resulting from an action or omission—related to medication management or care tasks performed at home by the caregiver, regardless of whether it caused harm to the care recipient [26]. Care recipients were classified as high complexity if their Barthel index was less than or equal to 60.

The design and reporting of this study followed key elements outlined in the CONSORT 2025 statement [27], providing a transparent and

comprehensive account of the VR training (Supplementary material S1). In addition, outcome evaluation was structured according to the Kirkpatrick Model [28,29], which includes four levels of analysis: participants' reaction to the training (Level 1), learning or knowledge acquisition (Level 2), behavioral changes in caregiving practices (Level 3), and impact on care quality or safety outcomes (Level 4).

3.1. Ethics

The study was approved by the Ethics Committee of Sant Joan d'Alacant University Hospital on February 1st, 2023 (project code: 22/080) and registered on ClinicalTrials.gov (ref. NCT05885334, March 31st, 2023; <https://clinicaltrials.gov/study/NCT05885334>). Data were collected anonymously to ensure participant confidentiality. All procedures complied with Spanish national regulations for research involving human participants and adhered to the principles of the Declaration of Helsinki (last updated October 2024).

3.2. Patient or public involvement

No patients directly participated in this study. Participants in the experimental group contributed insights that informed the future refinement of the educational materials.

3.3. Changes to trial protocol

Participation among caregivers of highly dependent patients was limited, extending the study timeline. Local logistics also reduced recruiting sites from nine to six and lowered the planned sample size, though statistical power was preserved.

3.4. Trial setting

The study was conducted in collaboration with healthcare institutions located in three Spanish regions: Andalusia, Madrid, and the Valencian Community.

3.5. Study period

The study was conducted from August 2023 to July 2025.

3.6. Participants

Informal caregivers of patients with multiple chronic conditions and responsible for providing care for at least six months per year. Inclusion required caregivers to have sufficient availability to complete baseline assessments, attend the intervention, and participate in follow-up. Caregivers were excluded if they had formal healthcare training or provided care to institutionalized patients for three or more months per year. Caregivers who were already familiar with the use of VR were also excluded.

3.7. Sample size

Based on previous studies reporting a 70% baseline prevalence of medication errors among informal caregivers [7], a sample size of 84 participants per group was determined to detect a 10-percentage-point reduction. This estimation was calculated using a two-tailed significance level of 5%, a statistical power of 80%, and an expected dropout rate of 15%.

3.8. Recruitment

Participants were recruited via multiple channels: routine clinic visits (when accompanying patients), caregiver training programs, and collaborations with associations/foundations. Interested individuals

were placed on preliminary lists; recruiters at this stage were unaware of future group allocation. Eligible caregivers were then randomly assigned to intervention or control using simple randomization (drawing lots). Allocation was conducted independently by an outcome assessor to maintain concealment and minimize selection bias. Group assignment was revealed only after informed consent, and participants were not told about the other group's activities to reduce bias.

3.9. Blinding

A team member independent of outcome assessment generated the simple randomization sequence with centralized allocation concealment. Given the intervention, participants and trainers could not be blinded, but data analysts remained blinded to group assignment.

3.10. Materials

We developed a library of immersive VR scenarios. Common preventable caregiving errors were identified from prior studies and semi-structured interviews with 12 clinicians (each > 5 years in home-hospitalization or medium-stay care). From these inputs, priority tasks were defined and scenarios scripted (e.g., hygiene, mobilization, medication management), then reviewed by an expert panel for relevance and clinical accuracy. A specialized vendor produced 18 scenarios for Meta Quest 2, enabling interactive, experiential practice of typical home-care situations. In addition, standard video scenarios depicting frequent errors were recorded for outcome assessment. All materials were piloted to verify usability and suitability.

3.11. VR onboarding

Before the intervention, participants in the experimental group completed a brief standardized orientation on headset fit, basic navigation, and the gesture set for the caregiving tasks to reduce motion-tracking issues and streamline onboarding.

3.12. Interventions

Participants in the control group received routine guidance and educational materials depicting the same home-care activities practiced in VR (Supplementary Materials S2). The experimental group completed a structured immersive VR session in designated labs. Training was delivered in VR-ready rooms cleared of obstacles and overseen by a facilitator; headset covers were used and equipment disinfected between users, and contraindications were screened beforehand. First-person simulations of common home-care tasks (e.g., hygiene, mobilization, medication administration) enabled active practice in a controlled environment. Scenarios were tailored to typical home contexts and could be repeated as needed. VR-trained staff provided technical support, navigation help, and safety oversight throughout.

The flow of the study is presented in Fig. 1.

3.13. Measures

This study employed four primary outcome measures aligned with the Kirkpatrick Model of training evaluation: Satisfaction (level 1), knowledge acquisition observable (level 2), behavior change (level 3), and signals of impact on care quality (level 4).

Data was collected at three time points—baseline, immediately post-intervention, and three-months after (follow-up), allowing for the analysis of both retrospective and prospective trends.

3.13.1. Satisfaction (Level 1)

Perceived usefulness, clarity, relevance, and overall value of the intervention were assessed using an ad hoc questionnaire adapted from tools based on the Technology Acceptance Model [30] (TAM).

3.13.2. Knowledge acquisition observable (level 2)

The ability of participants in the experimental group to detect potential errors in standard video recordings of simulated caregiving situations (featuring actors intentionally making common caregiving mistakes) was assessed. It was applied only to the experimental group.

3.13.3. Behavior change (level 3)

Participants' caregiving error frequency was analyzed. Participants in both the control and experimental groups also reported the types and consequences of errors they believed they had made. Additionally, perceived enhancement in caregiving skills was assessed using an adapted version of the IEXPAC questionnaire [31] for informal caregivers [32].

Possible restrictions imposed by the physical and social environment where caregiving took place were assessed using the six-item COM-B questionnaire [33], with responses rated on a scale from 0 (completely limited) to 10 (not limited at all).

3.13.4. Signals of impact on care quality (level 4)

The emotional burden experienced by caregivers was assessed using the short version of the Zarit Burden Interview (7 items [34], ZBI-7).

3.14. Data analysis

Satisfaction was summarized with descriptive statistics (frequencies and % positive at follow-up). Between-group differences in satisfaction by sex and patient complexity were explored using Fisher's exact test. For knowledge acquisition, within-group change (baseline vs follow-up) in the number of errors detected was assessed with the Wilcoxon signed-rank test due to non-normality. In addition, participants were classified as improved, unchanged, or worsened according to the direction of change in the number of errors detected. A two-sided sign test was used to assess directional change, with the null hypothesis that, among participants showing any change, improvement and worsening were equally likely. Participants with no change were treated as ties and excluded from the sign test.

To assess self-reported caregiving/medication errors, generalized

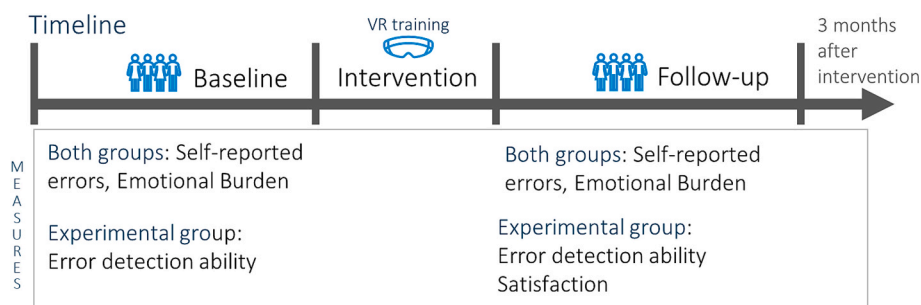


Fig. 1.

linear mixed models (GLMM) were used to evaluate the effect of phase (baseline vs follow-up, representing the effect of the intervention) and group, including their interaction, accounting for repeated measures within participants. Additional GLMMs were used to assess the effects of gender and patient complexity. A within-group Wilcoxon signed-rank test was also computed for the experimental arm as supplementary analysis.

For COM-B and ZBI-7, linear mixed-effects models (LMM) were used as the primary analyses to evaluate changes from baseline to follow-up, accounting for within-subject correlations. In addition, Wilcoxon signed-rank tests were performed for COM-B subscales and total scores, and for ZBI-7 subgroup contrasts. Emotional-burden categories (low, moderate, high) were summarized descriptively in both groups, and within the experimental group shifts in these categories from baseline to follow-up were examined with the McNemar–Bowker test as a secondary exploratory analysis (two-sided $\alpha = 0.05$).

4. Results

A total of 140 informal caregivers were enrolled in the study, of whom 70 were allocated to the experimental group and 70 to the control group. Fig. 2 provides the CONSORT 2025 flow diagram [27].

The mean age of participants was 44.6 years (SD = 19.9, with 109/140 (77.9%) identifying as female. Most caregivers were family members 115/140 (82.1%) primarily caring for older adults with chronic conditions such as Alzheimer's disease, reduced mobility, heart disease or Parkinson (Table 1).

4.1. Level 1: Satisfaction

Most participants (>90%) reported high satisfaction with the VR program: 98.6% (69/70) felt the time was worthwhile and the simulation engaging; ~91% (64/70) understood objectives, found it applicable to their work, and ~71% (50/70) adapted well. About 91% (64/70) said the simulation reflected daily caregiving and improved task performance (Supplementary Material S3). No sex differences appeared in any satisfaction item (Supplementary Material S4). Caregivers of high-complexity patients were more likely to see their work reflected ($p = 0.0036$) (Supplementary Material S5).

4.2. Level 2: Knowledge acquisition

In the experimental group, 46/70 (65.7%) participants increased the number of errors detected at follow-up, 11/70 (15.7%) showed no change, and 13/70 (18.6%) detected fewer errors than at baseline. Among participants who changed, improvement was significantly more frequent than worsening (two-sided sign test, $p < 0.0001$) (Table 2). Consistently, the mean number of errors identified increased from 5.41 (SD 1.94) at baseline to 6.64 (SD 2.56) at follow-up, an absolute mean gain of 1.23 errors per participant ($P = 0.0001$; Supplementary material S6). Among participants in the experimental group, women correctly identified more care-related errors than men at follow-up ($p = 0.0456$) (Supplementary material S7). Patient complexity did not influence the number of correctly identified errors ($p < 0.1730$) (Supplementary material S8).

4.3. Level 3: Observable behavior change

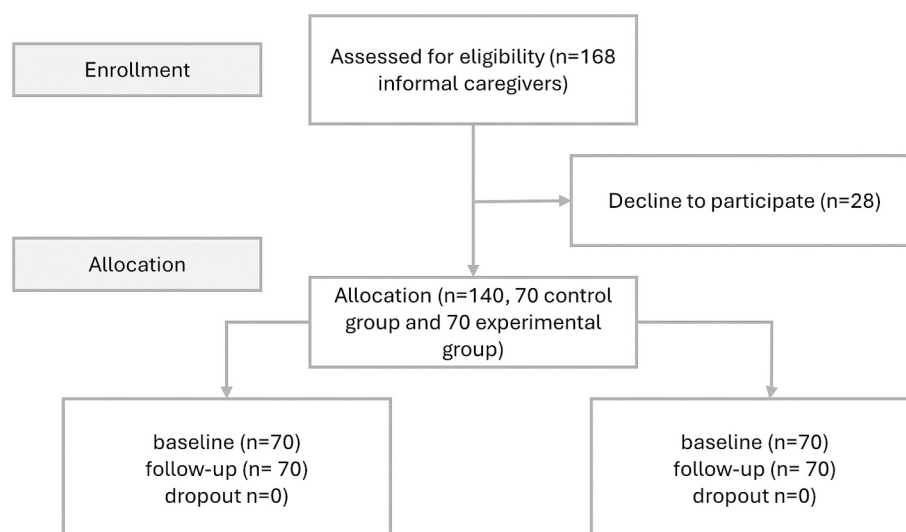
The experimental group significantly reduced the number of self-reported errors, from 62 ($M = 0.89$) at baseline to 23 ($M = 0.33$) at follow-up, while the control group increased their self-reported errors. This differential change between groups was statistically significant ($p < 0.0001$) (Table 3).

By gender, male errors remained at 4, whereas female errors declined from 58 to 19; GLMM showed no main effects of gender ($p = 0.9735$) (Supplementary S9). No main effects of complexity ($p = 0.2499$) were identified (Supplementary S10).

Among VR participants who reported at least one medication error, the distribution of error types changed from baseline ($N = 32$) to follow-up ($N = 19$). Timing errors decreased from 34.4% at baseline to 21.5% at follow-up (Supplementary material S11). The proportion with no consequences increased from 50.0% ($N = 14$) to 83.3% at follow-up ($N = 6$) (Supplementary material S12). From baseline to follow-up, the COM-B total score increased from 6.38 to 7.43 ($p = 0.0017$) (Table 4).

4.4. Level 4: Signals of impact on care quality

In the experimental group, emotional burden decreased from baseline to follow-up (Table 5). However, LLM showed a significant effect of group, but no significant effect of phase or group \times phase interaction (Supplementary material S14).



Based on: Hopewell S, Chan AW, Collins GS, Hróbjartsson A, Moher D, Schulz KF, et al. CONSORT 2025 Statement: updated guideline for reporting randomized trials. *BMJ*. 2025; 388:e081123.

Fig. 2.

Table 1
Samples description.

	Control Group	Experimental Group	Total
<i>Gender</i>			
Male	18 (25.7)	12 (17.1)	30 (21.4)
Female	51 (72.9)	58 (82.9)	109 (77.9)
Non-binary	1 (1.4)	0 (0.0)	1 (0.7)
Age (Mean, SD)	55.8 (11.6)	33.4 (20.2)	44.6 (19.9)
Relative	66 (94.3)	49 (70.0)	115 (82.1)
Living together	52 (74.3)	27 (38.6)	79 (56.4)
Years of caregiving experience (Mean, SD)	7.2 (11.1)	6.7 (8.3)	6.9 (9.8)
Hours per day (Mean, SD)	13.3 (8.8)	10.3 (9.5)	11.9 (9.3)
Formal training in caregiving	10 (14.3)	22 (31.4)	32 (22.9)
Spanish as primary language	67 (95.7)	65 (92.9)	132 (94.3)
<i>Patient sex</i>			
Male	35 (50.0)	35 (50.0)	70 (50.0)
Female	35 (50.0)	35 (50.0)	70 (50.0)
Patient age (Mean, SD)	75.6 (14.1)	66.9 (27.7)	71.5 (22.0)
Patient Barthel (0–100 points)	49.4 (35.6)	58.4 (26.7)	53.4 (32.1)
High complexity patient (Barthel ≤ 60)	41 (58.6)	41 (58.6)	82 (58.6)
Medications taken daily (Mean, SD)	7.7 (4.0)	4.6 (4.4)	6.3 (4.4)
Uses medication dispensing devices	41 (58.6)	32 (45.7)	73 (52.1)
Degree of recognized dependency	23 (32.9)	35 (50.0)	58 (41.4)
<i>Principal diagnosis</i>			
Alzheimer's disease	10 (14.3)	11 (15.7)	21 (15.0)
Cancer	11 (15.7)	7 (10.0)	18 (12.9)
Reduced mobility	6 (8.6)	7 (10.0)	13 (9.3)
Heart disease	10 (14.3)	2 (2.9)	12 (8.6)
Parkinson's disease	0 (0.0)	11 (15.7)	11 (7.9)
Chronic Obstructive Pulmonary Disease (COPD)	6 (8.6)	3 (4.3)	9 (6.4)
Dementia	5 (7.1)	2 (2.9)	7 (5.0)
Diabetes	2 (2.9)	4 (5.7)	6 (4.3)
Rare disease	2 (2.9)	2 (2.9)	4 (2.9)
Others	18 (25.7)	21 (30.0)	39 (27.9)
Total	70 (100)	70 (100)	140 (100)

Table 2
Change in the number of errors from baseline to follow-up in the experimental group (N = 70).

Change in error detection	N	%	p-value*
Detected more errors after the intervention	46	65.7	<0.0001
Detected the same number of errors after the intervention	11	15.7	
Detected fewer errors after the intervention	13	18.6	
Total	70	100.0	

*Two-sided sign test.

In contrast, COM-B scores showed significant effects of group and phase, as well as a significant group × phase interaction in LMM analyses, supporting a positive effect of the intervention on behavioral determinants (Supplementary material S15).

In the experimental group, caregiver empowerment ratings were high across items (means 8.2–8.9 on a 1–10 scale) in the follow-up (Supplementary material S16).

Table 3
Total number of self-reported caregiving errors by participants in control and experimental groups.

	Control Group (N = 70)		Experimental group (N = 70)	
	Baseline	Follow-up	Baseline	Follow-up
Medication errors	31	23	27	7
Care provision errors	15	54	35	16
Total errors self-reported	46	77	62	23
Effect (Intercept)	Estimate	Std. Error	z-value	P-value
Group (between)	−0.7662	0.2651	−2.890	0.0039
Phase (within)	−1.4071	0.3849	−3.656	0.0003
Group × Phase	−0.5269	0.1892	−2.785	0.0053
	1.6510	0.3265	5.057	<0.0001

*Generalized linear mixed-effects model.

Table 4
COM-B outcomes at baseline and follow-up in the experimental group (N = 70).

	Baseline Mean (SD)	Follow-up Mean (SD)	P-Value*
Average score for Opportunity	6.29 (1.90)	7.03 (2.37)	0.0325
Average score for Motivation	5.81 (1.24)	7.10 (2.21)	0.1078
Average score for Capability	7.25 (1.84)	7.93 (1.85)	0.0925
Total score (range 1–10)	6.38 (1.02)	7.43 (1.71)	0.0017

Table 5
Emotional burden of participants in control and experimental groups.

Emotional burden	Control group		Experimental group	
	Baseline (N, %)	Follow-up (N, %)	Baseline (N, %)	Follow-up (N, %)
Never	10 (14.3)	8 (11.4)	6 (8.6)	20 (28.6)
Rarely	8 (11.4)	38 (54.3)	9 (12.9)	18 (25.7)
Sometimes	24 (34.3)	11 (15.7)	40 (57.1)	24 (34.3)
Often	23 (32.9)	11 (15.7)	13 (18.6)	8 (11.4)
Almost always	5 (7.1)	2 (2.9)	2 (2.9)	0 (0.0)
Total	70 (100)	70 (100)	70 (100.0)	70 (100.0)

5. Discussion

5.1. Main findings

Participants reported high satisfaction and engagement; most found the training useful, realistic, and applicable. The intervention improved knowledge and performance and reduced emotional burden. Caregivers—including older adults with limited digital experience—adapted with minimal support, indicating feasibility for broader use.

The simulation integrated technical, contextual, and emotional fidelity to support comprehensive learning [35]; these elements were deliberately engineered, though high fidelity alone does not ensure impact.

Few programs employ immersive tech in realistic caregiver scenarios. Here, trainees rated VR highly, improved performance, and showed tangible practice changes, including lower burden. The structured program targeting caregiving/medication errors yielded gains across outcomes.

These findings reinforce the idea that immersive simulations, when designed according to user-centered principles and grounded in safety science, can effectively address the training gap among non-professional caregivers [16]. By focusing on frequent, preventable home-care errors, our high-fidelity scenarios offer a replicable, scalable approach to improving safety and quality at home.

5.2. Placing the findings in context

Immersive technologies have gained traction across various health-care domains [36–40]. VR-based training has demonstrated effectiveness in improving outcomes for patients with chronic conditions [41–44]. The VR capacity to simulate complex care environments in safe, controlled settings has been shown to enhance learning, motivation, and user engagement [16].

Consistent with previous research [36], our study confirms that VR training can enhance caregivers' ability to perform care tasks more safely. Immersive simulations support the acquisition of competence [45]. Additional elements, including haptic feedback, motion tracking, and immersive narratives, are linked to enhanced engagement [46], which is critical for durable behavior change. Subsequent research could isolate the effects of these attributes to determine their contribution to training efficacy.

Nonetheless, VR training outside clinical settings remains rare. Barriers include low digital literacy, reluctance to use technology—especially among older adults—and the absence of standardized protocols for home caregiving's cognitive, spatial, and emotional demands [38–40]. Our study shows feasibility in older caregivers, with high satisfaction, good adaptation, and measurable competence gains. Cost-effectiveness still requires evaluation, but results support broader rollout. Moreover, the potential for medium-term cost-efficiency—particularly through brief, high-intensity training delivered to large caregiver populations—is aligned with prior research [37,39]. As deinstitutionalization policies continue to expand across Europe and reliance on unpaid caregivers intensifies, our findings underscore the strategic value of immersive technologies in enabling safer, more competent, and more dignified care in domestic settings.

5.3. Practical implications

By 2021, unpaid caregiving represented an estimated 2%–10% of GDP across Europe and is likely to grow as health systems struggle to meet rising long-term care needs [2]. With limited institutional capacity, reliance on home-based care—often by untrained caregivers—will increase, underscoring the need for cost-effective strategies to protect patient safety at home. Many countries now offer allowances, legal recognition, and social protections, yet assuring quality in private settings remains challenging due to limited oversight. Structured training—such as the VR program evaluated here—can help close this gap by strengthening caregivers' competencies. VR supports knowledge retention, decision-making, and meaningful learning, and brief, high-intensity sessions at scale may be cost-effective. Adoption, however, requires upfront investment and basic technical skills. Our scenario library and standardized procedure indicate the feasibility of VR in this context.

5.4. Strengths

It is among the few randomized controlled trials evaluating VR training for informal caregivers under real-world conditions. The user-centered intervention used scenarios built from documented caregiver errors and expert input, ensuring contextual realism. Training content was designed, piloted, and refined for clinical accuracy and usability. Overall, the study offers a replicable, structured approach that provides accessible, effective tools for untrained caregivers, improving the safety and quality of home care.

5.5. Limitations

Because the intervention could not be blinded for participants or trainers, performance/expectancy effects are possible. Behavior change and error reduction were self-reported, so recall and social-desirability bias may persist despite anonymization and objective video checks.

The three-month follow-up limits inferences long-term skill retention and sustained change. Intervention participants were younger and differed in caregiving context/prior training, potentially favoring VR adaptability and inflating effects. We lacked objective clinical endpoints (e.g., home-verified medication discrepancies) and health-service use, limiting claims about patient outcomes and system impact. Finally, delivery in a supervised Spanish healthcare setting with trained staff may not mirror conditions for large-scale community deployment.

5.6. Future research

Future work should test skill transfer to real-life caregiving through direct observation in homes. Comparative studies must evaluate cost-effectiveness and knowledge retention versus online courses, video simulations, or in-person sessions. Usability studies are also needed on cybersickness and visual fatigue, especially in older adults. Research should examine psychological effects (e.g., self-efficacy, perceived competence). Finally, explore integrating biofeedback, voice interaction, and mixed reality to further personalized training.

6. Conclusion

Brief immersive VR training on common home-care scenarios improved recognition of unsafe practices, reduced self-reported caregiving/medication errors, and lowered emotional burden in informal caregivers. It also appeared to empower users and provide positive experience, with high satisfaction and gains in safe-care conditions and perceived capability. VR could complement caregiver support via brief, scalable sessions. Larger pragmatic trials and economic evaluations are needed.

7. Declaration of generative AI use

During the preparation of this work, the authors used OpenAI's ChatGPT to assist with translation from Spanish to English, language editing, and consistency checks to flag potential discrepancies between the text and the numerical data, tables, and figures. After using this tool, the authors independently verified all content and take full responsibility for the published article.

Data Availability

Aggregated data and study materials (e.g., questionnaires, scenario scripts) are available from the corresponding author upon reasonable request.

CRedit authorship contribution statement

Eva Gil-Hernández: Writing – review & editing, Visualization, Methodology, Investigation, Data curation. **Irene Carrillo:** Writing – review & editing, Methodology, Investigation. **Clara Pérez-Esteve:** Writing – review & editing, Visualization, Formal analysis, Data curation. **Almudena Arroyo:** Writing – review & editing, Investigation. **Mercedes Guilabert:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Investigation. **Purificación Ballester:** Writing – review & editing, Methodology, Investigation. **José Joaquín Mira:** Writing – review & editing, Writing – original draft, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

Funding

This study was conducted in the framework of the RealityCare activities funded by MCIN (Ministry of Science and Innovation)/AEI (State Research Agency; DOI 10.13039/501100011033; TED 2021-130383B-I00), and by the European Union “NextGenerationEU”/PRTR

(Recovery, Transformation and Resilience Plan). JJM enjoyed an intensified research activity contract from the Carlos III Health Institute (ISCIII) (reference INT22/00012). During the execution of this study and composition of this manuscript, EGH received funding through a Pre-doctoral Fellowship for Research Training in Health from Carlos III Health Institute (ISCIII), supported by the European Union NextGenerationEU and the Recovery, Transformation, and Resilience Plan from the Spanish Government (reference FI22/00277).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the volunteers from the collaborating institutions who facilitated and supported the conduct of this study, and the participants in both groups for their time and their commitment to patient safety at home. We also thanks Alejandro Rodríguez for managing the scenarios and providing support for the VR training.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijmedinf.2026.106444>.

References

- V. Zigante, Informal care in Europe: exploring formalisation, availability and quality, brussels: european commission, directorate-general for employment, Soc. Affairs Inclus. (2018), <https://doi.org/10.2767/78836>.
- European Parliament, Informal and long-term care in the EU – the role of gender and the impact on women’s economic independence, Brussels: Directorate-General for Internal Policies, Policy Department for Citizens’ Rights and Constitutional Affairs, Report No. IPOL_STU(2021)695476 (2021). doi: 10.2861/395998.
- K.A. Look, J.A. Stone, Medication management activities performed by informal caregivers of older adults, *Res. Social Adm. Pharm.* 14 (2018) 418–426.
- R. Schulz, J. Eden (Eds.), *Families Caring for an Aging America*, National Academies Press, Washington, DC, 2016.
- A. Parand, S. Garfield, C. Vincent, B.D. Franklin, Carers’ medication administration errors in the domiciliary setting: a systematic review, *PLoS One* 12 (2017) e0188631.
- I. Zygouri, F. Cowdell, A. Ploumis, et al., Gendered experiences of providing informal care for older people: a systematic review and thematic synthesis, *BMC Health Serv. Res.* 21 (2021) 730, <https://doi.org/10.1186/s12913-021-06736-2>.
- E. Gil-Hernández, P. Ballester, M. Guilabert, M.T. de Castro, Á. Cobos-Vargas, P. Pérez-Pérez, I. Carrillo, A.M. Fernández-Navascués, J.J. Mira, Enhancing safe medication use in home care: insights from informal caregivers, *Front. Med. (Lausanne)* 11 (2024) 1494771, <https://doi.org/10.3389/fmed.2024.1494771>.
- P. Ballester, C. Pérez-Esteve, A. Sánchez-García, E. Gil-Hernández, M. Guilabert, J. J. Mira, Understanding the roles of over 65-year-old male and female carers: a comparative analysis of informal caregiving, *Geriatrics (basel)* 10 (2025) 75, <https://doi.org/10.3390/geriatrics10030075>.
- R. Schulz, S.R. Beach, S.J. Czaja, L.M. Martire, J.K. Monin, Family caregiving for older adults, *Annu. Rev. Psychol.* 71 (2020) 635–659, <https://doi.org/10.1146/annurev-psych-010419-050754>.
- J.K. Parmar, T. L’Heureux, S. Anderson, et al., Optimizing the integration of family caregivers in the delivery of person-centered care: evaluation of an educational program for the healthcare workforce, *BMC Health Serv. Res.* 22 (2022) 364.
- R. Daynes-Kearney, S. Gallagher, Online support groups for family caregivers: scoping review, *J. Med. Internet Res.* 25 (2023) e46858, <https://doi.org/10.2196/46858>.
- D. Bressington, R. Gray, K. Barclay, C. Overy, I. Blackberry, Type and efficacy of online training for informal carers: a narrative review, *Front. Public Health* 13 (2025) 1603020, <https://doi.org/10.3389/fpubh.2025.1603020>.
- Y. Hur, R.L. Hickman Jr, Use and impact of simulation in family caregiver education: a systematic review, *West. J. Nurs. Res.* 46 (2024) 159–173, <https://doi.org/10.1177/01939459231218956>.
- C. Chen, S. Ding, J. Wang, Digital health for aging populations, *Nat. Med.* 29 (2023) 1623–1630, <https://doi.org/10.1038/s41591-023-02391-8>.
- Z.B. Kingir Baycar, M. Sancar, P. Ay, R. Demirtunc, C. Nuhoglu, C.H. Misirli, B. Okuyan, Short-term outcome of personalized theory-based pharmaceutical care service on the medication administration problems of family caregivers: a pre-post intervention study, *Int. J. Clin. Pharm.* (2024), <https://doi.org/10.1007/s11096-024-01823-w>.
- J. Wang, Q. Li, J. Cui, S. Tu, Z. Deng, R. Yang, Y. Wang, Effectiveness of virtual reality on the caregiving competence and empathy of caregivers for elderly with chronic diseases: a systematic review and meta-analysis, *J. Nurs. Manag.* 2023 (2023) 5449955, <https://doi.org/10.1155/2023/5449955>.
- M. Brydon, J. Kimber, M. Sponagle, J. MacLaine, J. Avery, L. Pyke, R. Gilbert, Virtual reality as a tool for eliciting empathetic behaviour in carers: an integrative review, *J. Med. Imaging Radiat. Sci.* 52 (2021) 466–477, <https://doi.org/10.1016/j.jmir.2021.04.005>.
- C.E. Miyawaki, A. McClellan, M.E. Kunik, et al., Feasibility and efficacy of life review delivered by virtually-trained family caregivers, *J. Alzheimers Dis.* 95 (2023) 581–593, <https://doi.org/10.3233/JAD-230371>.
- C. Rodríguez Villegas, Simuladores de entrenamiento basados en realidad virtual: aplicaciones en medicina, *Rev. Fac. Med. Hum.* 19 (2019) 137–143.
- A.A. Navarro-Newball, Entrenamiento en competencias usando realidad virtual, *Rev. Virtualidad Educ. Cienc.* 8 (2017) 75–93.
- D. Palesy, G. Forrest, M.E. Crowley, Education guidelines, frameworks and resources for building virtual care capacity: an integrative review, *J. Telemed. Telecare* 29 (2023) 153–164, <https://doi.org/10.1177/1357633X221149230>.
- P.B. Shani, N.J. Mathews, The rationale and evidence for virtual methods of training for caregivers: a narrative review, *J. Nurs. Res. Pract.* 4 (2020) 1–6.
- P. Budhrani-Shani, N. Mathews, The rationale and evidence for virtual methods of training for caregivers: a narrative review, *J. Nurs. Res. Pract.* 4 (2020) 1–6, [https://doi.org/10.37532/jnpr.2020.4\(2\).1-6](https://doi.org/10.37532/jnpr.2020.4(2).1-6).
- RealityCare, RealityCare project website (2019–2025). Available at: <https://realitycare.es> (Accessed 4 Oct 2025).
- N. Abed, B. Knezevic, P. Sousa, S. Tella, E. Srulovici, R. Strametz, J.J. Mira, Defining informal caregivers by their characteristics, safety roles and training needs in Europe, *Sci. Rep.* 15 (2025) 24375.
- M. Lisby, L.P. Nielsen, B. Brock, J. Mainz, How are medication errors defined? a systematic literature review of definitions and characteristics, *Int. J. Qual. Health Care* 22 (2010) 507–518.
- S. Hopewell, A.W. Chan, G.S. Collins, et al., CONSORT 2025 explanation and elaboration: updated guideline for reporting randomised trials, *BMJ* 389 (2025) e081124.
- D.L. Kirkpatrick, Evaluating training programs: the four levels, Berrett-Koehler, San Francisco, 1994.
- J.D. Kirkpatrick, W.K. Kirkpatrick, Kirkpatrick’s four Levels of Training Evaluation, ATD Press, Alexandria, VA, 2016.
- A.T. Lee, et al., Understanding psychosocial barriers to healthcare technology adoption: a review of TAM and UTAUT frameworks, *Healthcare (Basel)* 13 (2025) 250.
- R. Nuño, P. Fernández, J.J. Mira, N. Toro, J.C. Contel, M. Guilabert, O. Solas, Desarrollo de IEMAC, un instrumento para la evaluación de modelos asistenciales ante la cronicidad, *Gac. Sanit.* 27 (2013) 128–134.
- M. Guilabert, et al., The measure of the family caregivers’ experience, *Int. J. Environ. Res. Public Health* 15 (2018) 2040.
- C. Keyworth, T. Epton, J. Goldthorpe, R. Calam, C.J. Armitage, Acceptability, reliability, and validity of a brief measure of capabilities, opportunities, and motivations (COM-B), *Br. J. Health Psychol.* 25 (2020) 474–501.
- A.M. Gort, et al., Escala de Zarit reducida en cuidados paliativos, *Med. Clin. (Barc.)* 124 (2005) 651–653.
- A.K. Lefor, K. Harada, H. Kawahira, M. Mitsuishi, The effect of simulator fidelity on procedure skill training: a literature review, *Int. J. Med. Educ.* 11 (2020) 97–106.
- P. Parekh, et al., Systematic review and meta-analysis of augmented reality in medicine, retail, and games, *Vis. Comput. Ind. Biomed. Art* 3 (2020) 21.
- A. Herur-Raman, et al., Next-generation simulation—integrating extended reality into medical education, *Front. Virtual Real.* 2 (2021) 693399.
- M. Schreiter, J. Hennrich, A.L. Wolf, T. Eymann, The influence of previous experience on virtual reality adoption in medical rehabilitation and overcoming knowledge gaps among health care professionals: qualitative interview study, *J. Med. Internet Res.* 27 (2025) e26249.
- M. Kafes, Y.Y. Ileri, Current status of virtual reality research in healthcare: thematic and bibliometric analysis, *Front. Virtual Real.* 6 (2025), <https://doi.org/10.3389/frvir.2025.1411075>.
- S.M. Alexander, V. Friedman, P.M. Rerkpattanapipat, W.A. Hiatt, J.S. Heneghan, R. Hubal, Y.Z. Lee, Adapting novel augmented reality devices for patient simulations in medical education, *Cureus* 16 (2024) e66209, <https://doi.org/10.7759/cureus.66209>.
- H.S. Lee, Y.J. Park, S.W. Park, The effects of virtual reality training on function in chronic stroke patients: a systematic review and meta-analysis, *Biomed Res. Int.* 2019 (2019) 7595639, <https://doi.org/10.1155/2019/7595639>.
- Y. Chen, A. Fu, et al., Effects of virtual reality rehabilitation on upper limb function and balance in patients with Parkinson’s disease: a meta-analysis, *Medicine (Baltimore)* 101 (2022) e28717.
- G. Li Pira, B. Aquilini, A. Davoli, S. Grandi, C. Ruini, The use of virtual reality interventions to promote positive mental health: systematic literature review, *JMIR Ment Health* 10 (2023) e44998, <https://doi.org/10.2196/44998>.
- M.E. Ding, H. Traiba, H.R. Perez, Virtual reality interventions and chronic pain: scoping review, *J. Med. Internet Res.* 27 (2025) e59922.
- S. Licata, L. Deroma, C. Fresco, A. Di Chiara, Virtual reality simulation training improves health professionals’ knowledge in STEMI critical care, *Eur. J. Public Health* 33 (2023) ckad160.806, <https://doi.org/10.1093/eurpub/ckad160.806>.
- C.J. Clay, J.R. Budde, A.Q. Hoang, A. Gushchyan, An evaluation of the effectiveness of immersive virtual reality training in non-specialized medical procedures for caregivers and students: a brief literature review, *Front. Virtual Real.* 5 (2024).

Glossary

COM-B: Capability, Opportunity, Motivation – Behavior (behavior change model)
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: CONSORT extension for digital health interventions
COPD: Chronic obstructive pulmonary disease
df: Degrees of freedom
GDP: Gross domestic product
GLMM: Generalized linear mixed-effects model
IEXPAC: Instrument for Evaluation of Chronic Patient Experience (adapted for caregivers)
LTC: Long-term care

LMM: Linear mixed-effects model
M: Mean
N: Sample size (number of observations/participants)
p: p-value (probability value)
RCT: Randomized controlled trial
SD: Standard deviation
Sign test: Non-parametric test for paired directional change
TAM: Technology Acceptance Model
VR: Virtual reality
ZBI-7: 7-item Zarit Burden Interview (short form)