Journal of Healthcare Quality Research xxx (xxxx) 101226

Journal of Healthcare Quality Research

www.elsevier.es/jhqr

VSI: ADVANCING PATIENT SAFETY

# Exploring sex differences in immersive safety training for home-based care among informal caregivers: A randomized comparative study

**P. Ballester[a,1], E. Gil-Hernández[b,1], C. Pérez-Esteve[b], A. Arroyo[c], M. Guilabert[d], I. Carrillo[d], J.J. Mira[b,d,*], J. Vitaller[b,e]**

[a] *Facultad de Farmacia y Nutrición, Universidad Católica de Murcia, Murcia, Spain*
[b] *Grupo de Investigación ATENEA, Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana (FISABIO), Alicante, 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] *Departamento de Psicología de la Salud, Universidad Miguel Hernández de Elche, Elche, Spain*
[e] *Inspección Servicios Sanitarios, Dirección Territorial de Sanidad en Alicante, Elche, Spain*

Received 4 May 2026; accepted 30 June 2026

**KEYWORDS**
Informal caregivers;
Patient safety;
Home-based care;
Virtual reality;
Self-perceived errors;
Sex differences

**Abstract**
*Background:* Informal caregivers are key to home-based care safety.
*Objective:* To evaluate whether the effectiveness of immersive training aimed at reducing home-based care errors differs according to caregiver sex.
*Methods:* A randomized comparative study with pre–post assessment involving 141 informal caregivers compared virtual reality (VR; $n = 70$) and 360° video training (360VT; $n = 71$). Data were collected at baseline and three months after the intervention. The primary outcome was the number of self-reported home-based care errors. Data were analysed using a mixed-effects Poisson regression model.
*Results:* The mean number of self-reported errors decreased from 0.91 to 0.46 per caregiver after training. Exploratory subgroup analyses suggested that pre–post changes may have differed according to training modality and caregiver sex, with larger estimated reductions among women in the VR group (IRR = 0.29, 95%CI: 0.14–0.59; $p = 0.0005$) and among men in the 360VT group (IRR = 0.19, 95%CI: 0.06–0.60; $p = 0.0048$). The phase × modality × sex interaction was statistically significant (IRR = 13.88, 95%CI: 1.95–98.75; $p = 0.0087$). Satisfaction was high overall, although women in the VR group reported lower ease of participation (adjusted $p = 0.0373$).

* Corresponding author.
*E-mail address:* jose.mira@umh.es (J.J. Mira).
[1] Joint first authors.

https://doi.org/10.1016/j.jhqr.2026.101226


Please cite this article as: P. Ballester, E. Gil-Hernández, C. Pérez-Esteve et al., Exploring sex differences in immersive safety training for home-based care among informal caregivers: A randomized comparative study, Journal of Healthcare Quality Research, https://doi.org/10.1016/j.jhqr.2026.101226

*Conclusions:* Immersive training may contribute to improving home-based care safety among informal caregivers. These findings also highlight the importance of appropriate onboarding, usability adaptations, and further evaluation in larger and more diverse caregiver populations.


**PALABRAS CLAVE**
Cuidadores informales;
Seguridad del paciente;
Atención domiciliaria;
Realidad virtual;
Errores autodeclarados;
Diferencias por sexo

**Diferencias por sexo en una intervención formativa inmersiva sobre seguridad en el cuidado domiciliario dirigida a cuidadores informales: estudio comparativo aleatorizado**

**Resumen**
*Antecedentes:* Los cuidadores informales son clave para la seguridad en la atención domiciliaria.
*Objetivo:* Evaluar si los resultados de una formación inmersiva para reducir errores en los cuidados difieren según el sexo del cuidador.
*Métodos:* Estudio comparativo aleatorizado con evaluación pre-post que comparó entrenamiento con realidad virtual (RV; n = 70) y mediante vídeo 360° (360VT; n = 71). Los datos se recogieron antes y tres meses después de la intervención. La variable principal fue el número de errores en el cuidado domiciliario autodeclarados. Los datos se analizaron mediante un modelo de regresión de Poisson de efectos mixtos.
*Resultados:* El número medio de errores autodeclarados por cuidador disminuyó de 0,91 a 0,46 tras la formación. Los análisis exploratorios por subgrupos mostraron variaciones en los cambios pre-post según la modalidad y el sexo, con reducciones estimadas mayores entre las mujeres del grupo de RV (IRR = 0,29; IC95%: 0,14–0,59; p = 0,0005) y entre los hombres del grupo 360VT (IRR = 0,19; IC95%: 0,06–0,60; p = 0,0048). La interacción fase × modalidad × sexo fue estadísticamente significativa (IRR = 13,88; IC95%: 1,95–98,75; p = 0,0087). La satisfacción fue elevada en conjunto, aunque las mujeres del grupo de RV comunicaron una menor facilidad para participar (p ajustado = 0,0373).
*Conclusiones:* La formación inmersiva podría contribuir a mejorar la seguridad de los cuidados en el domicilio, aunque su implementación requiere una introducción adecuada a la tecnología y nuevos estudios con diferentes perfiles de cuidadores.


## Introduction

Informal caregiving is an essential source of support for people living with chronic conditions, multimorbidity, or functional dependence.[1] In Spain, 10.6% of the population provides unpaid care.[2] This contribution allows many care recipients to remain at home and delay or avoid institutionalization.

Despite their relevance, informal caregivers often remain insufficiently recognised by health and social welfare services. Research on integrated care shows that their needs are not always assessed systematically and that they are not consistently incorporated into care continuity, even though they sustain a large share of everyday care and are key actors in quality and safety.[3]

## Safety of the care recipient

Home is not a risk-free care environment.[4] Unintentional errors may occur in medication administration, hygiene, mobilization, aspiration prevention, monitoring of warning signs, or use of devices.[5] Recent studies emphasize risk anticipation and prevention as essential components of safe home-based care.[6] These incidents represent an important patient safety problem and position home care as a priority setting for preventive interventions. Their frequency also appears to be higher than previously assumed, which reinforces the need for specific training to create safer home environments.[6]

## Training for safety

Improving caregiver training is a growing priority. However, developing effective training interventions remains challenging. Caregivers often face time constraints and show wide variability in health literacy, digital skills, and confidence in using technology.[7–9] Existing interventions are heterogeneous and not always adapted to real home-based care contexts, highlighting the need for more context-sensitive approaches.[10]

Immersive educational technologies offer a promising alternative. Virtual reality (VR) has been associated with significant improvements in knowledge, skills, satisfaction, and confidence.[11] Immersive 360° video training (360VT) is also increasingly used because it can improve attention, motivation, and satisfaction, while offering a more accessible and less costly alternative.[12]

A further question is whether the use and effects of immersive training differ by sex. Available evidence suggests that these differences are not linear.[13,14] Early performance in VR has been linked to younger age, male sex, and prior videogame experience, indicating that initial advantages may reflect technological familiarity rather than inherent learning differences.[14] Importantly, better initial performance does not necessarily translate into better final outcomes. Some studies show that women may present less favourable initial attitudes or lower early technical performance but demonstrate greater improvement with practice.[14,15] Overall, sex-related differences in immersive environments appear more likely to emerge in early adaptation, perceived ease of use, or physical comfort than in final learning outcomes.[16]

Although most evidence on sex-related differences in immersive learning comes from students, healthcare professionals, or simulation-based procedural training, this question may be particularly relevant for informal caregivers. Caregiver populations are heterogeneous in age, digital literacy, previous exposure to immersive technologies, and confidence in using digital tools.[17,18] Future studies should examine not only whether immersive technologies improve safety-related skills, but also whether they are acceptable, usable, and adaptable for caregivers with different levels of age, digital literacy, confidence, and previous technological experience.[19] Given their potential to provide realistic, repeatable, and relatively low-cost training for practical home-care tasks,[20] VR and 360VT require evaluation in real-world caregiver populations to identify the support and adaptations needed for effective implementation.

However, evidence comparing different immersive training modalities in informal caregivers remains scarce, and little is known about whether caregiver sex influences training effectiveness or user experience in home care safety interventions. Therefore, this study aimed to examine whether women and men show different post-intervention improvement patterns according to digital training modality.

## Methods

### Study design

A randomized comparative study with pre–post assessment was conducted to compare two immersive training modalities, VR and 360VT, within a project focused on improving home-based care safety. Both technologies provide immersive learning experiences, although they differ substantially in the degree of interactivity they offer. Data were collected at baseline (pre-intervention) and three months after the intervention (post-intervention).

### Hypothesis

Given the limited evidence in informal caregivers, no directional superiority of one modality for either sex was assumed. It was hypothesized that changes in self-reported errors and user experience could differ according to the interaction between sex and training modality.

### Setting

The study was carried out in collaboration with healthcare institutions from three Spanish regions, Andalusia, Madrid, and the Valencian Community, between January 2024 and July 2025.

### Recruitment and allocation

Informal caregivers were recruited through healthcare visits, caregiver training activities, and collaborations with associations and foundations. This strategy enabled access to caregivers with diverse profiles and real-world home-based care experience.

Recruitment and allocation were implemented by the collaborating centres, with no involvement of the central research team in participant identification, invitation, recruitment, or assignment to study conditions. Potential participants were identified and invited by the collaborating centres. After eligibility had been confirmed, informed consent had been obtained, and baseline assessment had been completed, the local centre assigned each participant to VR or 360VT according to its predefined random allocation procedure.

Because the participating centres differed in their organisation, workflow, and recruitment routes, a single centralized allocation sequence was not imposed. Instead, each collaborating centre applied a locally prespecified random allocation procedure adapted to its operational context. Depending on the centre, allocation was implemented through randomly assigned participant entry numbers, individual random assignment as caregivers entered the study, or randomly assigned recruitment days for each condition. A local allocation log was used to monitor recruitment progress and ensure that the prespecified number of participants was reached in each condition. These procedures were intended to preserve random allocation while avoiding assignment based on participant characteristics, staff preference, convenience, or expected response to the intervention. The central research team did not have access to the lists of people contacted or to those who declined participation. Allocation was known only to the local staff responsible for organising the intervention sessions. The statistical analysis was conducted using an anonymised database by researchers who were not involved in recruitment, allocation, or intervention delivery. Randomization was not stratified by sex; therefore, the sex-by-modality analyses were prespecified as exploratory analyses.

In this study, an informal caregiver was defined as an unpaid person, usually a family member, partner, friend, or relative, who regularly provides help, supervision, or direct care for daily living, basic care, or medication management.[21]

Eligibility criteria required participants to provide home-based care on a stable basis and to be available to complete baseline assessment and the training session. Individuals with formal healthcare training, those caring for institutionalized persons for three or more months per year, and those with previous experience using VR were excluded in order to reduce bias related to prior technological familiarity.

## Materials and procedure

The immersive materials were designed to reproduce frequent and potentially unsafe home-based care situations, based on preventable errors identified in previous studies[22,23] which included interviews, focus groups and pilots.

In this study, self-reported errors refer to caregivers' retrospective recognition of preventable unsafe actions or omissions in medication management and daily care tasks, such as unsafe combinations of essential medication; unsafe mobilization with risk of falls; inappropriate feeding practices in people with dysphagia, including inadequate food texture or positioning, with risk of choking or aspiration; insufficient prevention of pressure injuries or community-acquired infections due to inadequate hygiene practice; delayed recognition of warning signs of deterioration; or incorrect use of home-based care devices such as oxygen therapy, feeding tubes, catheters, or mobility aids.

In the VR arm, a library of interactive scenarios for Meta Quest 2 was developed. In the 360VT arm, immersive observational scenes were produced based on the same domains and learning objectives. A total of 18 training scenarios were available, covering the main domains of home-based care safety, including medication management, mobilization and transfers, hygiene, feeding and dysphagia-related risks, recognition of warning signs, and use of home-based care devices. Each participant completed between three and four scenarios, selected according to the care tasks they regularly performed at home.

Both training modalities shared the same learning objectives and safety domains, but differed in format and level of interactivity. In the VR arm, participants first received a standardized introductory session of approximately 15 min on headset use, navigation, interaction commands, and safety instructions. They then completed a supervised practice session of approximately 45 min using Meta Quest 2 headsets. The VR scenarios were fully interactive: participants navigated through a digital home environment, interacted with a virtual care recipient, and performed care-related actions while following on-screen instructions.

In the 360VT arm, participants also received initial instructions on how to access and use the materials. The videos were available on their own mobile phone or tablet and could be viewed as often as needed before, after, or when safely preparing care tasks. Participants could move the device to observe the scene from different angles and follow the explanations provided in each scenario. Unlike VR, 360VT was mainly observational and self-paced, although it addressed the same care situations and preventive messages.

VR sessions were supervised by trained staff to assist participants, ensure correct use of the equipment, and prevent accidents. For the 360VT group, support was available after the initial session through a telephone support service for technical or practical questions. Facilitators had previous experience with the technologies and received standardized training using a common protocol to minimize differences in the information provided to participants. Safety and hygiene procedures were applied, including cleaning of headset pads, use of obstacle-free spaces, screening and instructions for dizziness or cybersickness, and secure access to the digital materials.

Both interventions covered identical educational content and differed only in the level of interactivity provided by the immersive technology.

## Study variables

The main outcome was the number of self-reported home-based care errors in the previous three months. Self-reported errors were used as a pragmatic indicator of perceived safety-related caregiving behaviour based on previous studies.[5,6,24] Operationally, a caregiver error was defined as any preventable event, by action or omission, related to medication management or care tasks performed at home, regardless of whether harm occurred to the care recipient. This measure was collected at two time points: baseline and three months after the intervention. At baseline, participants reported the number of errors perceived during the three months prior to the intervention, whereas at follow-up they reported errors perceived during the three months after the intervention. This outcome does not capture objectively observed errors or adverse events. Moreover, it must be considered that training may alter caregivers' awareness of unsafe practices and their threshold for identifying an event as an error.

Participants were instructed to report situations that they considered to be care errors, regardless of whether harm had occurred.

The secondary outcome was satisfaction with the intervention. It included perceived value of time invested, attractiveness of the simulation, understanding of scenario objectives, applicability to caregiving tasks, ease of participation, adaptation to the tool, realism of the content, and perceived improvement in task performance. The satisfaction questionnaire was adapted from instruments inspired by the Technology Acceptance Model.[25]

## Statistical analysis

To explore whether changes in the number of self-reported errors over time varied by sex and training modality, a mixed-effects Poisson regression model was fitted with assessment phase (pre vs post), intervention type (VR vs 360VT), sex and age as fixed effects, and participant as a random intercept. The outcome was based on the number of errors reported for the previous three months, collected at baseline (pre-intervention) and three months after the intervention. Results were expressed as beta coefficients, incidence rate ratios (IRR), 95% confidence intervals, and $p$ values.

Given the exploratory nature of the sex-by-modality analyses and the presence of baseline imbalances, sensitivity

analyses were conducted to assess the stability of the findings. The full model was compared with a reduced model excluding the phase × modality × sex interaction using AIC and a likelihood ratio test. An Age × modality interaction was also tested to examine whether age modified the effect of training modality. Finally, models additionally adjusted for baseline errors and post−pre change scores adjusted for baseline errors were fitted to assess whether baseline differences or regression to the mean could explain the observed pattern.

Satisfaction was summarized using frequencies and proportions of positive responses. Comparisons were made across the subgroups defined by intervention and sex using the Fisher-Freeman-Halton exact test. For items with significant global differences, pairwise Fisher's exact tests with Bonferroni correction were performed.

### Ethical considerations

The study was approved by the Ethics Committee of Hospital Universitario Sant Joan d'Alacant (February 1, 2023, project code: 22/79) and registered in ClinicalTrials (NCT05885347). Data collection procedures complied with applicable Spanish regulations and were conducted in accordance with the ethical principles of the Declaration of Helsinki, as revised in October 2024.

Generative AI (ChatGPT, OpenAI, version 5.3) was used exclusively as a language-support tool, specifically to assist with the English translation of the manuscript and to improve clarity and readability. The study design, scientific content, data analyses, interpretation of findings, and conclusions were entirely developed by the authors.

## Results

A total of 141 informal caregivers participated, of whom 112 were female (79.4%) and 29 were male (20.6%). Overall, 71 participants (50.4%) were assigned to the 360VT condition and 70 (49.6%) to the VR condition. Baseline characteristics of participants are presented in Table 1.

### Self-reported caregiving errors

No significant differences were observed between groups at baseline (IRR = 1.35, 95%CI: 0.68−2.68; $p = 0.3922$), although a trend towards higher error rates was found in men (IRR = 2.24, 95%CI: 0.90−5.57; $p = 0.0813$). Age was not significantly associated with self-reported errors (IRR = 1.00, 95%CI: 0.98−1.02; $p = 0.8119$).

Interaction terms suggested that changes in self-reported errors over time may have varied according to modality and sex. VR was associated with a greater reduction in self-reported errors after training among women (IRR = 0.29, 95%CI: 0.14−0.59; $p = 0.0005$ for the phase × modality interaction), whereas men showed a greater reduction in the 360VT modality (IRR = 0.19, 95%CI: 0.06−0.60; $p = 0.0048$ for the phase × sex interaction). The statistically significant three-way interaction suggested a possible differential pattern of change across sex and modality subgroups (IRR = 13.88, 95%CI: 1.95−98.75; $p = 0.0087$). The observed pattern was compatible with differential responses by sex and modality, with larger reductions among women in the VR group and among men in the 360VT group. Detailed regression estimates are presented in Table 2.

These findings were consistent with the observed pre−post changes across subgroups, which showed significant reductions in errors among men in the 360VT group and women in the VR group, but not in the other groups (Table 3).

Additional sensitivity analyses were conducted to examine whether the observed sex-by-modality pattern was influenced by model specification, age imbalance, or baseline differences in self-reported errors (Table 4). The full model showed better fit than the model excluding the phase × modality × sex interaction, while adding an age × modality interaction did not improve model fit. Models accounting for baseline errors and post−pre change scores showed a similar direction of findings. Taken together, these analyses indicate that the pattern was not solely attributable to baseline imbalances or regression to the mean.

In addition to effectiveness outcomes, participants' satisfaction with the immersive training was also evaluated.

### Satisfaction with the immersive training

Subjective evaluation of the intervention was positive across groups (Table 5). For most items, more than 90% of participants gave positive responses, regardless of intervention type or sex. However, global differences emerged for two dimensions. The first was understanding of the objective of each situation through signs and audio instructions ($p = 0.0475$), for which females exposed to VR showed a somewhat lower proportion of positive responses. The second was ease of participation in the study ($p = 0.0086$), also rated less favourably by females in the VR group. In addition, the item related to having been trained in something applicable to caregiving showed a trend towards significance ($p = 0.0722$) with the same pattern.

Pairwise comparisons showed that, for understanding of the objective of the situations, no difference remained significant after Bonferroni correction (Table 6). In contrast, for ease of participation, a significant difference remained between females in the 360VT group and females in the VR group (adjusted $p = 0.0373$), with better ratings in the 360VT group (Table 7). Therefore, there were indications of a less favourable user experience among females who used VR.

## Discussion

These findings indicate that immersive training was associated with post-intervention improvement, reflected in a reduction in self-reported errors, although this effect was not homogeneous across subgroups. The significant three-way interaction between phase, modality, and sex suggests that the impact of the intervention depended on both training format and participant sex, although future studies should examine these relationships in greater depth.

Female caregivers exposed to VR showed a clear reduction in self-reported errors, whereas among male caregivers the most pronounced improvement was observed in the

**Table 1** Sample description.

| | 360VT ($n = 71$) | RV ($n = 70$) | Total ($n = 141$) |
|---|---|---|---|
| *Sex* | | | |
| Male | 17 (23.9) | 12 (17.1) | 29 (20.6) |
| Female | 54 (76.1) | 58 (82.9) | 112 (79.4) |
| *Age (mean, SD)* | 23.4 (5.7) | 33.4 (20.2) | 28.3 (15.6) |
| *Hours of care per day (mean, SD)* | 8.8 (9.3) | 10.3 (9.5) | 9.6 (9.4) |
| *Different medications taken daily by the care recipient (mean, SD)* | 4.4 (3.4) | 4.6 (4.4) | 4.5 (3.9) |
| *Relative to care recipient* | 51 (71.8) | 49 (70.0) | 100 (70.9) |
| *Living with the care recipient* | 20 (28.2) | 27 (38.6) | 47 (33.3) |
| *Formal training related to caregiving* | | | |
| None | 61 (85.9) | 48 (68.6) | 109 (77.3) |
| Up to 20 h | 4 (5.6) | 5 (7.1) | 9 (6.4) |
| From 20 to 40 h | 3 (4.2) | 9 (12.9) | 12 (8.5) |
| More than 40 h | 3 (4.2) | 8 (11.4) | 11 (7.8) |
| *Use of a medication dispensing device* | | | |
| Yes | 37 (52.1) | 32 (45.7) | 69 (48.9) |
| No | 34 (47.9) | 38 (54.3) | 72 (51.1) |
| *Care tasks performed*[*] | | | |
| Preparing meals | 57 (80.3) | 46 (65.7) | 103 (73.0) |
| Modifying food texture | 17 (23.9) | 11 (15.7) | 28 (19.9) |
| Feeding | 25 (35.2) | 16 (22.9) | 41 (29.1) |
| Hygiene (shaving, tooth brushing) | 35 (49.3) | 26 (37.1) | 61 (43.3) |
| Bathing or showering | 38 (53.5) | 34 (48.6) | 72 (51.1) |
| Transfers (bed-chair/chair-toilet/chair-armchair) | 38 (53.5) | 30 (42.9) | 68 (48.2) |
| Wound care | 37 (52.1) | 15 (21.4) | 52 (36.9) |

Values are presented as *n* (%) for categorical variables and mean (SD) for continuous variables unless otherwise indicated.
[*] Multiple responses allowed.

**Table 2** Mixed-effects model for self-reported errors.

| Fixed effects | $\beta$ (SE) | IRR | 95%CI IRR | *p*-Value[*] |
|---|---|---|---|---|
| *Main effects* | | | | |
| Intercept | −0.95 (0.36) | 0.39 | 0.19–0.80 | 0.0084 |
| Phase (post vs pre) | −0.05 (0.22) | 0.95 | 0.62–1.46 | 0.8253 |
| Modality (VR vs 360VT) | 0.30 (0.35) | 1.35 | 0.68–2.68 | 0.3922 |
| Sex (male vs female) | 0.81 (0.46) | 2.24 | 0.90–5.57 | 0.0813 |
| Age (years) | −0.002 (0.01) | 1.00 | 0.98–1.02 | 0.8119 |
| *Interaction terms* | | | | |
| Phase × modality | −1.23 (0.35) | 0.29 | 0.14–0.59 | 0.0005 |
| Phase × sex | −1.65 (0.58) | 0.19 | 0.06–0.60 | 0.0048 |
| Modality × sex | −1.81 (0.84) | 0.16 | 0.03–0.83 | 0.0307 |
| Phase × modality × sex | 2.63 (1.00) | 13.88 | 1.95–98.75 | 0.0087 |
| *Random effects* | | | | |
| Participant (intercept), SD = 1.22 | | | | |

[*] Mixed-effects Poisson regression model.

360VT group. These results should not be interpreted as evidence of a globally superior modality, but rather as an indication that responses to immersive training may differ by subgroup. The apparent greater reduction among men in the 360VT group, for example, may partly reflect higher baseline error counts in that subgroup, whereas the low baseline level among men in the VR group suggests a possible floor effect. Given the small and unbalanced subgroup

**Table 3** Pre–post changes in the number of self-reported errors by sex and intervention.

| Group | Pre Mean (SD) | Post Mean (SD) | Mean change Mean (SD) | *p*-Value* |
|---|---|---|---|---|
| Male 360VT ($n = 17$) | 1.29 (1.36) | 0.24 (0.56) | −1.05 (1.52) | 0.0119 |
| Male VR ($n = 12$) | 0.33 (0.78) | 0.25 (0.45) | −0.08 (0.79) | 0.8501 |
| Female 360VT ($n = 54$) | 0.81 (1.57) | 0.78 (1.72) | −0.04 (1.84) | 0.4205 |
| Female VR ($n = 58$) | 1.00 (1.81) | 0.28 (0.62) | −0.72 (1.79) | 0.0023 |

* Wilcoxon signed-rank test.

**Table 4** Sensitivity analyses assessing the stability of the exploratory subgroup findings.

| Sensitivity analysis | Purpose | Model specification | Main finding | Interpretation |
|---|---|---|---|---|
| Reduced interaction model | To assess the contribution of the phase × modality × sex interaction | Model excluding the three-way interaction | AIC increased from 622.11 to 626.62; LRT $p = 0.0107$ | The full model showed better fit |
| Age effect modification | To assess whether age modified the association between modality and outcome | Main model plus age × modality interaction | AIC = 624.11; $p = 1.0000$ | There was no statistical evidence that age modified the modality effect |
| Baseline-adjusted model | To assess whether baseline differences in self-reported errors explained the findings | Main model additionally adjusted for baseline errors | Direction of subgroup effects remained consistent | Baseline differences did not fully account for the observed exploratory pattern |
| Change-score model | To assess baseline dependency and regression to the mean | Linear model of post–pre change scores adjusted for baseline errors | VR effect: $\beta = -0.56$; $p = 0.0136$ | Findings were consistent with the primary analysis and were not solely explained by regression to the mean |

sizes, and the baseline differences in age and self-reported errors, several sensitivity analyses were performed to examine the stability of the findings. These analyses showed that removing the phase × modality × sex interaction worsened model fit, whereas adding an age × modality interaction did not improve fit.

However, this pattern is consistent with previous research showing that sex-related differences in immersive environments are complex and not indicative of a stable advantage for either group.[11,13] Evidence suggests that initial differences, often favouring men, may reflect prior exposure to videogames or 3D technologies rather than inherent learning capacity.[13,26,27] Moreover, better initial performance does not necessarily translate into better outcomes, as some studies report steeper learning curves among women despite less favourable starting points.[14] Overall, such differences appear more likely to emerge in early adaptation, usability, or comfort, rather than in final learning outcomes.[28]

The role of individual characteristics in immersive learning is particularly relevant in procedural training contexts.[29] The present intervention aimed not only to transmit knowledge but also to influence safety-related caregiving behaviour. The reduction observed among women in the VR group may reflect the added value of interactivity, realistic exposure to risk situations, and active engagement in decision-making. On the other hand, these findings should not be taken as evidence that 360VT is intrinsically more effective for men, but rather that less interactive yet realistic formats may also support learning, particularly when they are accessible and easy to use.

Subjective experience showed relevant nuances. Although overall satisfaction was high across all subgroups, female caregivers exposed to VR rated ease of participation less favourably and, to a lesser extent, understanding of scenario objectives. Notably, this less favourable experience coexisted with a significant reduction in self-reported errors. This suggests that VR may be effective even when initial usability is perceived as more demanding. It also reinforces the idea that sex-related differences may emerge more clearly in early interaction with the technology rather than in outcomes.[16] These findings reinforce the importance of considering usability alongside effectiveness when implementing immersive educational interventions for informal caregivers.

The reduction observed among female caregivers exposed to VR is also consistent with previous evidence

**Table 5** Satisfaction with the immersive training by sex and modality.

| Item | 360VT ($n = 71$) | | VR ($n = 70$) | | *p*-Value* |
|---|---|---|---|---|---|
| | Male | Female | Male | Female | |
| *Was the time invested in the simulation worthwhile?* | | | | | |
| Yes | 16 (94.1) | 52 (96.3) | 12 (100) | 57 (98.3) | 0.5571 |
| No | 1 (5.9) | 2 (3.7) | 0 (0.0) | 1 (1.7) | |
| *Did you find the simulation attractive?* | | | | | |
| Yes | 16 (94.1) | 50 (92.6) | 12 (100) | 57 (98.3) | 0.3980 |
| No | 1 (5.9) | 4 (7.4) | 0 (0.0) | 1 (1.7) | |
| *Did you understand the objective of each scenario through signs and audio instructions?* | | | | | |
| Yes | 17 (100) | 54 (100) | 12 (100) | 52 (89.7) | 0.0475 |
| No | 0 (0.0) | 0 (0.0) | 0 (0.0) | 6 (10.3) | |
| *Did the training provide skills applicable to caregiving?* | | | | | |
| Yes | 17 (100) | 54 (100) | 11 (91.7) | 53 (91.4) | 0.0722 |
| No | 0 (0.0) | 0 (0.0) | 1 (8.3) | 5 (8.6) | |
| *Was participation in the study easy?* | | | | | |
| Yes | 16 (94.1) | 54 (100) | 10 (83.3) | 50 (86.2) | 0.0086 |
| No | 1 (5.9) | 0 (0.0) | 2 (16.7) | 8 (13.8) | |
| *Were you able to adapt to the tool?* | | | | | |
| Yes | 10 (58.8) | 37 (68.5) | 7 (58.3) | 43 (74.1) | 0.5117 |
| No | 7 (41.2) | 17 (31.5) | 5 (41.7) | 15 (25.9) | |
| *Did the simulation reflect your daily caregiving experience?* | | | | | |
| Yes | 16 (94.1) | 53 (98.1) | 10 (83.3) | 54 (93.1) | 0.1501 |
| No | 1 (5.9) | 1 (1.9) | 2 (16.7) | 4 (6.9) | |
| *Did the simulation improve your performance in the trained task?* | | | | | |
| Yes | 17 (100) | 52 (96.3) | 11 (91.7) | 54 (93.1) | 0.5817 |
| No | 0 (0.0) | 2 (3.7) | 1 (8.3) | 4 (6.9) | |
| *Total* | 17 (100) | 54 (100) | 12 (100) | 58 (100) | |

* Differences across groups were assessed using the Fisher–Freeman–Halton exact test.

**Table 6** Pairwise comparisons for the item ''Did you understand the objective of each scenario through signs and audio instructions?'' (overall *p*-value = 0.0475).

| Group 1 | Group 2 | *p*-Value | Adjusted *p*-value (Bonferroni) |
|---|---|---|---|
| Male 360VT | Female 360VT | 1.0000 | 1.0000 |
| Male 360VT | Male VR | 1.0000 | 1.0000 |
| Male 360VT | Female VR | 0.3266 | 1.0000 |
| Female 360VT | Male VR | 1.0000 | 1.0000 |
| Female 360VT | Female VR | 0.0277 | 0.1663 |
| Male VR | Female VR | 0.5806 | 1.0000 |

**Table 7** Pairwise comparisons for the item ''Was participation in the study easy?'' (overall *p*-value = 0.0086).

| Group 1 | Group 2 | *p*-Value | Adjusted *p*-value (Bonferroni) |
|---|---|---|---|
| Male 360VT | Female 360VT | 0.2394 | 1.0000 |
| Male 360VT | Male VR | 0.5534 | 1.0000 |
| Male 360VT | Female VR | 0.6744 | 1.0000 |
| Female 360VT | Male VR | 0.0308 | 0.1846 |
| Female 360VT | Female VR | 0.0062 | 0.0373 |
| Male VR | Female VR | 0.6779 | 1.0000 |

on the added value of interactive immersive training. VR combines immersion with interactivity, which may enhance engagement, self-efficacy, and active learning.[11,12,16] Several mechanisms may explain these exploratory findings, although the present study was not designed to establish causal explanations. This may support transfer to caregiving behaviour and potentially safer practices at home. However, this interpretation should be cautious, as the outcome was based on self-report and does not directly measure objective safety improvements. Alternative explanations should also be considered. The observed reduction among women in the VR group may not be attributable solely to the interactive nature of VR. Baseline differences, regression to the mean, changes in awareness of what constitutes an error, social desirability,[30] and unmeasured differences in digital literacy[31] or previous exposure to immersive technologies may have contributed to the observed pattern.

From an applied perspective, these findings do not support sex-specific training content but do highlight the importance of implementation strategies. Brief onboarding, familiarization with the virtual environment, clear initial instructions, and ergonomic adjustments may help reduce early usability barriers, particularly among users with lower technological familiarity. These considerations are especially relevant in caregiver populations, where variability in age, digital literacy, and prior exposure to technology is high.

The results also have implications for patient safety in home-based care. Informal caregivers often perform complex tasks without formal training. The reduction in self-reported errors suggests that immersive training may improve awareness of unsafe situations and promote reflection on caregiving practices. However, self-reported errors should be considered an indirect indicator of safety. Changes in reporting may reflect behavioural improvement, increased awareness, or altered interpretation of errors. Future studies should combine self-reported measures with objective indicators, such as observed performance or follow-up assessments of specific behaviours.

This study has several strengths. It addresses an underexplored but highly relevant setting for patient safety. The intervention was based on realistic scenarios informed by evidence and expert input, supporting its contextual relevance. It also examines subgroup differences and includes user experience measures, enhancing its applicability to real-world implementation.

However, limitations should be acknowledged. The main outcome relied on self-reported errors, which may be subject to recall bias and social desirability. Subgroups were unbalanced, particularly for men, limiting statistical power and contributing to wide confidence intervals. Baseline differences may also have influenced the magnitude of improvement. The baseline age imbalance between groups, particularly the older and more heterogeneous age distribution in the VR arm, may have influenced usability, adaptation to the technology, and intervention effects despite adjustment for age. Additionally, data on videogame experience, digital literacy, spatial ability, or cybersickness were not collected, which could have helped interpret differences in user experience. The satisfaction measure was ad-hoc and not fully validated. Generalizability is limited by the Spanish setting, the predominance of female caregivers, and the exclusion of participants with previous VR experience. The findings may be most applicable to caregiver populations with limited prior exposure to immersive technologies and should not be directly extrapolated to more technologically experienced users.

Despite these limitations, this study has several important strengths, including its randomized comparative design, multicentre recruitment across three Spanish regions, the direct comparison of two immersive training modalities, and the longitudinal assessment of safety-related outcomes.

Future studies should confirm these findings in larger and more diverse caregiver populations, incorporating objective measures of caregiving performance and patient safety outcomes to better understand the role of immersive training in home care.

## Conclusions

Immersive training was associated with a reduction in self-reported errors among informal caregivers. An exploratory pattern suggested that changes may vary according to sex and training modality, but these subgroup findings should be interpreted cautiously because of small and unbalanced subgroup sizes, baseline differences, wide confidence intervals, and the self-reported nature of the outcome. Overall, the results suggest the potential value of immersive training for improving safety in informal home-based care, while highlighting the need for careful onboarding, usability adaptations, and longer-term evaluation using more robust safety outcomes.

## Funding

This study has been funded by the Instituto de Salud Carlos III (Spain), and co-funded by the European Union (ERDF, ''A way to make Europe''), reference PI22/00868. JJM enjoyed an intensified research activity contract granted by Instituto de Salud Carlos III (ISCIII) (reference INT22/00012). During the execution of this study and composition of this manuscript, EG-H received funding through a Predoctoral Fellowship for Research Training in Health from Instituto de Salud Carlos III (ISCIII), supported by the European Union NextGenerationEU and the Recovery, Transformation, and Resilience Plan from the Spanish Government (reference FI22/00277).

## Conflict of interest

The authors declare no conflict of interest.

## References

1. Bernini A, Natale F, Nedee A, Querin F. Informal long-term care in the context of demographic change and intergenerational fairness in the EU. Luxembourg: Publications Office of the European Union; 2026.
2. Ministerio de Sanidad. Encuesta de Salud de España 2023. Proyecto técnico y principales resultados. Madrid: Ministerio de Sanidad; 2025.
3. Ambugo EA, de Bruin SR, Masana L, MacInnes J, Mateu NC, Hagen TP, et al. A cross-European study of informal carers' needs in the

context of caring for older people, and their experiences with professionals working in integrated care settings. Int J Integr Care. 2021;21:2.
4. Shahrestanaki SK, Rafii F, Najafi Ghezeljeh T, Farahani MA, Majdabadi Kohne ZA. Patient safety in home health care: a grounded theory study. BMC Health Serv Res. 2023;23:467.
5. Parand A, Garfield S, Vincent C, Franklin BD. Carers' medication administration errors in the domiciliary setting: a systematic review. PLoS One. 2016;11, e0167204.
6. Gil-Hernández E, Ballester P, Guilabert M, Sánchez-García A, García-Torres D, Astier-Peña MP, et al. Enhancing safe medication use in home-based care: insights from informal caregivers. Front Med (Lausanne). 2024;11, 1494771.
7. Oh E, Moon S, Chung D, Choi R, Hong GRS. The moderating effect of care time on care-related characteristics and caregiver burden: differences between formal and informal caregivers of dependent older adults. Front Public Health. 2024;12, 1354263.
8. Moore C, Gallagher P, Dunne S. Health literacy, eHealth literacy and their association with burden, distress, and self-efficacy among cancer caregivers. Front Psychol. 2024;15, 1283227.
9. Chen Q, Zhang H, Yuan S, Liu W, Lyu T. Perceived training needs of the informal caregivers of older adults: a cross-sectional study. Healthcare (Basel). 2024;12:2369.
10. McGuigan K, Laurente G, Christie A, Noble H. Effectiveness of interventions for informal caregivers of people with end-stage chronic illness: a systematic review. Syst Rev. 2024;13:245.
11. Sung H, Kim M, Park J, Shin N, Han Y. Effectiveness of virtual reality in healthcare education: systematic review and meta-analysis. Sustainability. 2024;16:8520.
12. Blair C, Walsh C, Best P. Immersive 360° videos in health and social care education: a scoping review. BMC Med Educ. 2021;21:590.
13. Katz D, Hyers B, Patten E, Sarte D, Loo M, Burnett GW, et al. Relationship between demographic and social variables and performance in virtual reality among healthcare personnel: an observational study. BMC Med Educ. 2024;24:227.
14. Tebala GD, Duro F, Tató I, De Bonis Cristalli C, Properzi L, Cirocchi R, et al. The impact of gender, videogaming and music playing on robotic surgery simulation performance. Surg Laparosc Endosc Percutan Tech. 2026;36:e1420.
15. Mahling M, Wunderlich R, Steiner D, Gorgati E, Festl-Wietek T, Herrmann-Werner A. Virtual reality for emergency medicine training in medical school: prospective, large-cohort implementation study. J Med Internet Res. 2023;25, e43649.
16. Chang YC, Lo YT, Yang CC. Effects of interactivity, immersion, and physical discomfort on learning in VR nursing education. PLoS One. 2026;21, e0344586.
17. National Academies of Sciences, Engineering, and Medicine. Caregivers: diversity in demographics, capacities, and needs. In: Reducing the impact of dementia in America: a decadal survey of the behavioral and social sciences. Washington, DC: The National Academies Press; 2021. p. 109–36.
18. Lee S, Ory MG, Vollmer Dahlke D, Smith ML. Technology use among older adults and their caregivers: cross-sectional survey study. JMIR Aging. 2024;7, e50759.
19. Khan S, Webster S, Puxty J, Robertson M. Key challenges and barriers to digital literacy for older adults: scoping review. JMIR Aging. 2026;9, e80647.
20. Mira J, Pérez-Esteve C, Gil-Hernández E, Arroyo A, Aranaz J, Ballester P, et al. Improving home-based care safety among informal caregivers through immersive digital simulation: secondary analysis of three coordinated intervention studies. J Med Internet Res. 2026;28:e85141.
21. Abed N, Knezevic B, Sousa P, Tella S, Srulovici E, Strametz R, et al. Defining informal caregivers by their characteristics, safety roles and training needs in Europe. Sci Rep. 2025;15, 24375.
22. Mira J, Rios J, Gil-Hernández E, Abed N, Ribeiro-Neves V, Pérez-Esteve C, et al. Performance and perceptions of health care professionals using an immersive virtual reality tool for home-based care training: observational feasibility and acceptability study. JMIR Serious Games. 2025;13, e75104.
23. Abed N, Neves VR, Perez VM, Ortola DM, García VF, Knezevic B, et al. Consensus-based safety recommendations for informal home-based care: a mixed methods study. Res Nurs Health. 2026, http://dx.doi.org/10.1002/nur.70074.
24. González de Dios J, López-Pineda A, Mira-Perceval Juan G, Alcalá Minagorre PJ, Guilabert M, Pérez-Jover V, et al. Perceptions and attitudes of pediatricians and families with regard to pediatric medication errors at home. BMC Pediatrics. 2023;23:380.
25. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. 1989;13:319–40.
26. Grassini S, Laumann K. Are modern head-mounted displays sexist? A systematic review on gender differences in HMD-mediated virtual reality. Front Psychol. 2020;11:1604.
27. Feng J, Spence I, Pratt J. Playing an action video game reduces gender differences in spatial cognition. Psychol Sci. 2007;18:850–5.
28. van Dun C, van Kraaij A, Wegman J, Kuipers J, Aarts E, Janzen G. Sex differences and the role of gaming experience in spatial cognition performance in primary school children: an exploratory study. Brain Sci. 2021;11:886.
29. Duruaku F, Sims V, Jentsch FG. Individual differences influencing procedural learning outcomes in virtual reality: a case study on an exterior preflight inspection. Front Virtual Real. 2025;6, 1601562.
30. Vesely S, Klöckner CA. Social desirability in environmental psychology research: three meta-analyses. Front Psychol. 2020;11:1395.
31. Chang CY, Kuo HC, Du Z. The role of digital literacy in augmented, virtual, and mixed reality in popular science education: a review study and an educational framework development. Virtual Reality. 2023;27:2461–79.