Socially assistive robots for people with dementia: Systematic review and meta-analysis of feasibility, acceptability and the effect on cognition, neuropsychiatric symptoms and quality of life

Clare Yu, Andrew Sommerlad, Lena Sakure, Gill Livingston

UCL Division of Psychiatry, London, UK
Camden and Islington NHS Foundation Trust, London UK
UCL Division of Psychology and Language Sciences, London, UK

ARTICLE INFO

Keywords:
Dementia
Older adult
Robot
Neuropsychiatric
Depression
Quality of life

ABSTRACT

Background: There is increasing interest in using robots to support dementia care but little consensus on the evidence for their use. The aim of the study is to review evidence about feasibility, acceptability and clinical effectiveness of socially assistive robots used for people with dementia.

Method: We conducted a systematic review and meta-analysis. We searched MEDLINE, EMBASE, PsychINFO, CINHAL, IEEE Xplore Digital Library, and EI Engineering Village from inception to 04—02—2022 - included primary studies assessing feasibility, acceptability, or effectiveness of socially assistive robots for people with dementia. Two independent reviewers screened studies for eligibility, and assessed quality. Narrative synthesis prioritized higher quality studies, and random-effect meta-analyses compared robots with usual care (UC) or active control (AC) immediately after the intervention (short-term; ST) or long-term (LT) on cognition, neuropsychiatric symptoms, and quality of life.

Findings: 66 studies and four categories of robots were eligible: Companion robots (Pet and humanoid companion robots), telepresence communication robots, homecare assistive robots and multifunctional robots. PARO (companion robot seal) was feasible and acceptable but limited by its weight, cost, and sound. On meta-analysis, PARO had no ST or LT compared to UC or AC over 5—12 weeks on agitation (ST vs UC, 4 trials, 153 participants: pooled standardized mean difference (SMD) 0.25; −0.57 to 0.06; LT vs UC; 2 trials, 77 participants, SMD = −0.24; −0.94, 0.46), cognition (ST vs UC, 3 trials, 128 participants: SMD = −0.03; −0.32, 0.38), overall neuropsychiatric symptoms (ST vs UC, 3 trials, 169 participants: SMD = −0.01; −0.32, 0.29; ST vs AC, 2 trials, 145 participants: SMD = −0.02, −0.71, 0.85), apathy (ST vs AC, 2 trials, 81 participants: SMD = −0.14; 0.29, 0.58), depression (ST vs UC, 4 trials, 181 participants; SMD = 0.08; −0.52, 0.69; LT vs UC: 2 trials, 77 participants: SMD = −0.01; −0.75, 0.77), anxiety (ST vs UC: 2 trials, 104 participants, SMD = 0.24; −0.85, 1.33) and quality of life (ST vs UC, 2 trials, 127 participants: SMD = −0.05; −0.52, 0.42; ST vs AC: 2 trials, 159 participants, SMD = −0.36, −0.76, 0.05). Robotic animals, humanoid companion robots, telepresence robots and multifunctional robots were feasible and acceptable. However, humanoid companion robots have speech recognition problems, and telepresence robots and multifunctional robots were often difficult to use. There was mixed evidence about the feasibility of homecare robots. There was little evidence on any of these robots’ effectiveness.

Conclusion: Although robots were generally feasible and acceptable, there is no clear evidence that people with dementia derive benefit from robots for cognition, neuropsychiatric symptoms, or quality of life. We recommend that future research should use high quality designs to establish evidence of effectiveness.
1. Introduction

The number of people with dementia is projected to increase from around 57 million globally in 2015 to 152 million by 2050 (Nichols et al., 2022). It is expected by 2030, the global cost could grow to US$2 trillion, with detrimental implications for social and health-care systems (Wimo et al., 2017). Many countries in the world are also experiencing a shortage of dementia workers. It is expected by 2030 an additional 3.5 million care staff will be required in the United States alone (Prince et al., 2015), making provision of adequate social care for people with dementia a major public health concern. The COVID-19 pandemic has further led to staff shortages in health and social care and a reduction in the level of support provided (Society, 2020). Assistive robots may play an important role in this era by providing emotional support, companionship, and automated homecare support to people with dementia.

Different types of assistive robots have been developed. (Feil-Seifer and Mataric, 2005; Hegel et al., 2009) Physically assistive robots such as robotic wheelchairs support the user’s physical, and mobility needs but do not target other symptoms or socio-emotional needs. Socially assistive robots, however, are equipped with a social interface to enable interaction with the user and there are several subtypes. (Kachouie et al., 2014) For example, service robots such as homecare assistive robots provide supervision and monitoring to support daily activities at home. Robotic therapists provide training and therapies. (Ismail et al., 2017; Martin et al., 2013) Companion robots such as PARO, a seal-like robot, and cat and dog robots can interact with users through sounds and movements. Telepresence communication robots facilitate video chats between people living apart. PARO is by far the most extensively researched robot in dementia care. Several trials had reported benefits of PARO, such as improving quality of life (Joranson et al., 2016c), mood (In Soon and Hee Sun, 2018; Inoue et al., 2021b; Kelly et al., 2021; Takayanagi et al., 2014), pain (Joranson et al., 2016b), and agitation (Lane et al., 2016b; Pu et al., 2022). There are existing systematic reviews on this topic, but the published evidence for use of robots in dementia had mixed findings, likely due to mixed methodological approach and diverse study quality of these reviews (Hung et al., 2019b; Leng et al., 2019; Moyle et al., 2017a). One found in six randomised controlled trials (RCT) that pet robots had significant effects on agitation and depression, but not on cognition or quality of life. (Leng et al., 2019)

A scoping review examined benefits and barriers of using PARO, a socially assistive robot seal, in dementia care and suggested it has the potential to improve mood and neuropsychiatric symptoms, but that high cost, heavy staff workload, and infection concerns could inhibit its use in practice. (Hung et al., 2019b) Another review examined the feasibility of telepresence robots for people with dementia, concluding they have the potential to improve social connections but technical issues can inhibit their use. (Moyle et al., 2017a) However these reviews have been limited by not assessing the quality of the included papers (Moyle et al., 2017a) only considering health literature and papers in English language and excluding engineering databases, (Hung et al., 2019b) and not reporting if the effects of robots were greater than those of active controls, such as plush toys, or whether validated outcome measures were used in the studies. (Holthe et al., 2018; Hung et al., 2019b; Leng et al., 2019; Shu and Woo, 2021).

At present, robots are not routinely used in dementia care, but it is important to know about their feasibility, effectiveness and their implementation in real-world settings. Many factors could influence the potential widespread use of robots, such as environmental constraints, ease of use, and stakeholders’ attitudes towards the use of technology. (Holthe et al., 2018; Shu and Woo, 2021). Improving understanding of robots’ feasibility and effectiveness would inform research, clinical practice, and policy. A comprehensive review of feasibility of all types of socially assistive robots, and of effectiveness using only findings from validated effectiveness outcome instruments, would serve as a foundation for further research, improvement of robot design and to understand whether the current place of robots for people with dementia.

We aimed to conduct a comprehensive systematic review and meta-analyses that (1) examines feasibility and acceptability of robot use in dementia, (2) examines immediate and longer-term effectiveness compared to usual care and active controls of socially assistive robots used for people with dementia using validated outcome measures including cognition, neuropsychiatric symptoms, and quality of life. And (3) identifies facilitators and barriers to the use of robots for people with dementia. The study will shed light to the dementia care field by providing by summarizing all the existing knowledge in the field and provide a reliable guide for future clinical practice.

2. Method

This systematic review is reported in accordance with the Preferred Reporting Items for Systematic Review (PRISMA) 2019 (Moher et al., 2009) and registered on PROSPERO (CRD42020168239) on 19 March 2020.

2.1. Search strategy

We searched for articles indexed in six electronic databases (MEDLINE, EMBASE, PsycINFO, Cumulative Index of Nursing and Allied Health Literature (CINAHL), IEEE Xplore Digital Library and EI Engineering village) published from inception to 04 Feb 2022 and hand-searched the contents of existing systematic reviews to identify any other potentially relevant articles. There was no restriction on language. We searched for peer reviewed, primary studies on ‘Robots’ and ‘Dementia’, using Medical subject headings (Mesh) or where appropriate, the database-specific thesaurus equivalent. The search strategy is detailed in Appendix 1.

2.2. Study inclusion and exclusion criteria, and selection

There is no formal definition of a socially assistive robot, therefore we defined them as robots having a mechanical structure or design which enables interaction with a person with dementia to promote their well-being. (Bemelmans et al., 2012) We included primary quantitative and qualitative studies of socially assistive robots for people with dementia, where ≥70% participants had a dementia diagnosis or where results of people living with dementia were presented separately. We included studies assessing feasibility and acceptability from the perspective of stakeholders (people with dementia, caregivers, staff, dementia experts), or those assessing, using an instrument validated for people with dementia, effectiveness for cognition, neuropsychiatric symptoms (including overall neuropsychiatric symptoms, agitation, anxiety, depression, apathy), or quality of life. We excluded case studies and studies of several robot types where findings for specific robots could not be separated.

CY conducted the search. Two authors (CY and LS) then independently screened all titles and abstracts of the first 100 papers for inclusion. The inter-rater reliability was high (κ = 0.95). As no papers of the first 100 papers were excluded incorrectly by CY, she screened the remaining papers, and then retrieved the full text of any potentially eligible article. Both authors then independently reviewed all full text articles against the eligibility criteria. Any disagreement between the two authors was resolved through discussion with AS and GL.

2.3. Data extraction

CY extracted data on the characteristics of the included studies and all data were later cross-checked by LS. The extracted data includes (1) study design (RCT, non RCT, observational study, qualitative study), (2) participants’ characteristics (including mean age, sex, stage of dementia), (3) setting (place and country where the study was conducted), (4) description of the robotic intervention (e.g. Specific robot used, mode of delivery, duration of exposure), (5) comparator if used (e.g. usual care or
For qualitative studies, we additionally assessed the consideration of researchers and participants. Appendix 2 gives the details of the modiﬁed theoretical saturation and the reporting of the relationship between outcome assessments, and whether statistical analyses were appropriate. For non-randomized interventional studies, we added items assessing allocation concealment, the conduct of power calculation, sample size included studies. For RCT, we additionally assessed the adequacy of further appraisal items to improve assessment of the risk of bias for our studies, 4) qualitative studies, and 5) mixed-method studies. We added 2) non-randomized interventional studies, 3) quantitative descriptive this has five separate scales depending on the type of study for: 1) RCT, allocation concealment, the conduct of power calculation, sample size sufﬁciency, and the appropriateness of statistical analyses performed. For non-randomized interventional studies, we added items assessing the availability of a between- or within-subject control group, whether a power calculation was conducted, sample size sufﬁciency, blinding of outcome assessments, and whether statistical analyses were appropriate. For qualitative studies, we additionally assessed the consideration of theoretical saturation and the reporting of the relationship between researchers and participants. Appendix 2 gives the details of the modiﬁed MMAT tool used in this review.

CY and LS determined the study design and independently assessed study quality. Differences in study quality rating were discussed with AS and GL until we reached consensus. We then classiﬁed studies as high, moderate, or lower quality; appendix 3 details our criteria for this grading. Within each category of study quality, we prioritized ﬁndings from RCTs over ﬁndings from other types of study (non RCT, descriptive quantitative study, qualitative study).

2.5. Data synthesis and analysis

We reviewed the characteristics of robotic interventions of each included study and categorized them into different robotic groups. We operationalised deﬁnitions for each type of robot based on the major function of the robot. Four types of robots were identiﬁed, and we deﬁned them in the following way.

2.5.1. Companion robots

Robots that provide companionship and interacted with people with dementia to reduce social isolation/loneliness. We further divided this into pet companion robots (companion robots resembling any animal) and humanoid companion robots (companion robots resembling humans).

2.5.2. Telepresence communication robots

Robots that are teleoperated and allowed interactive chats between people with dementia and operator.

2.5.3. Multifunctional robots

Robots that provide multiple services, including any three of the followings: interactive talks, entertainment, therapeutic purposes, daily care activities, or telecommunication and other services.

2.5.4. Homecare assistive robots

Robots that solely support home-care activities. They include any one of the following functions: providing instructions for homecare tasks, managing medication, and providing reminders of daily activities. If the robots performed both a homecare supporting function and a non-homecare supporting function, we classiﬁed it as multifunctional robots instead.

We conducted a narrative synthesis for each robot type in three steps using Popay’s approach. (Popay et al., 2006) At step 1, CY and LS conducted a preliminary synthesis using tabulation as described above (See data extraction section). For each robot type, key ﬁndings related to feasibility and acceptability that fell under the pre-deﬁned data source of the Almere’s theoretical model of technology framework were tabulated (Appendix 4). This included attrition rate, attendance rate, and interview data reported from qualitative papers. Effectiveness ﬁndings for cognition, neuropsychiatric symptoms, and quality of life were also tabulated at this step. At step 2, CY scrutinised the tabulated key ﬁndings, to identify patterns/differences in acceptability, feasibility, and direction of effect across studies. Participants’ characteristics (e.g., dementia severity) and the context of the intervention (e.g., individual or group intervention format) were considered at this step. The identiﬁed patterns were then discussed with AS and GL until consensus was reached. At step 3, we assessed the robustness of the synthesis product we created at step 2, considering the risk of bias categorised above (see risk of bias assessment section). We used the evidence grading system detailed in Appendix 3 to provide strength of evidence of our synthesis, deﬁning studies as high, moderate, or lower quality.

We then meta-analysed studies of effectiveness which had similar interventions and outcome measures using RevMan 5.3 software. We conducted meta-analyses if there were two or more RCTs using the same type of robots. PARO was the only robotic intervention that met this criterion. We used random-effects models due to heterogeneity in study setting and samples. We calculated the standardized mean differences (SMD) for each outcome with 95% conﬁdence interval (CI) as studies used different outcome instruments. We used post-test score for SMD calculation. In cases in which studies only provided change-from-baseline score, we calculated the post-test score using baseline. If the studies did not report the post-test SD and no other data was provided to calculate such score, we used the baseline SD. For cluster RCTs, we used approximate analysis (Higgins et al., 2019) to reduce the size of each trial to its effective sample. We emailed the corresponding author for clariﬁcation if data were missing or unclear. One study (Moyle et al., 2017b) was excluded from the meta-analysis because we could not obtain required data. We used I2 to determine heterogeneity, with I2 > 75% indicating considerable heterogeneity. (Higgins et al., 2003). We planned to conduct sensitivity analyses assessing the effect of study quality on ﬁndings if there were two or more high quality RCTs but we could not do so as there were insufﬁcient high-quality studies.

3. Results

3.1. Search results

The PRISMA diagram shows the search results (Fig. 1). We identiﬁed 1211 records with 66 articles fulﬁlling inclusion criteria. Sixty-two were in English language, and there were one each in Korean (Song, 2009), French (de Sant’Anna et al., 2012) and two in German (Hielscher, 2015; Schramek et al., 2021).

3.2. Characteristics of included studies

There were 66 studies, including a total of 1750 people with dementia, 178 family caregivers, 232 staff and experts. Participants were from Europe, Australia, Canada New Zealand, US, India, Hong Kong, Korea, and Japan. Most participants with dementia were females living in a long-term care (LTC) setting. There were four categories of robots: companion robots (PARO, other pet robots, humanoid companion robots) (44 studies; 66%), multifunctional robots (15 studies; 23%), telepresence communication robots (3 studies; 5%) and homecare assistive robots (4 studies; 6%). Table 1 describes the main features of each robot.
3.3. Companion robots

There were 44 studies of companion robots, 29 used PARO (See Table 2 for high and moderate quality studies, and Appendix 5 for lower quality studies), 11 robotic dogs and cats (Table 3) and four humanoid companion robots (Table 4).

3.3.1. PARO

3.3.1.1. PARO’s feasibility and acceptability. PARO studies comprised 12 studies from Europe, seven studies from Australia, two studies from New Zealand, two from Korea, three from US, two from Japan and one from Canada. 24 studies examined feasibility and acceptability and 13 studied effectiveness Ten RCTs, (Joranson et al., 2015a; Joranson et al., 2016b; Koh and Kang, 2018; Liang et al., 2017; Moyle et al., 2013a, 2017b; Petersen et al., 2017; Pu et al., 2020; Valentí Soler et al., 2015), 11 other interventional studies/ observational studies (Bemelmans et al., 2016; de Sant’Anna et al., 2012; Demange et al., 2018; Jones et al., 2018; Joranson et al., 2016a; Kelly et al., 2021; Lane et al., 2016a; Robinson et al., 2013; Song, 2009; Takayanagi et al., 2014; Valentí Soler et al., 2015) and six qualitative studies reported feasibility and acceptability (Bemelmans et al., 2016, 2013; Hung et al., 2019a; Moyle et al., 2016, 2019; Pu et al., 2019) and two mixed method studies (Demange et al., 2019; Inoue et al., 2021a). There were five high quality, six moderate quality and 18 lower quality studies. They included 968 people with dementia, 76 family caregivers and 72 care staff. Studies measured feasibility and acceptability through drop-out and attendance rate, observations of engagement and enjoyment with attitudes towards PARO from surveys or interviews. PARO was generally acceptable to people with dementia.

All high-quality qualitative studies reported PARO had acceptable appearance, features, and was useful in improving mood, social connection, and building positive environment in the LTC facility (Hung et al., 2019a; Moyle et al., 2016; Moyle et al., 2019; Pu et al., 2019). The moderate quality RCT (Moyle et al., 2017b) reported a small dropout rate (3%) and high attendance rate (average number of attended sessions were 26/30).

The five interventional studies of moderate quality had similar findings. Low attrition rates (ranged from 0% to 10%) (Joranson et al., 2016b; Liang et al., 2017; Pu et al., 2020) and frequent attendance (85% participants attended 8 out of 12 sessions) (Song, 2009) were reported. However, adherence was lower when PARO was used at home (only 53% used PARO daily) (Liang et al., 2017). Lower quality studies reported similar findings (Appendix 5). Several limitations of PARO were reported in high and moderate quality studies, including it being viewed as childish (Bemelmans et al., 2016, 2013; Liang et al., 2017; Moyle et al., 2016, 2019), making disturbing noises when used (Liang et al., 2017; Moyle et al., 2016; Pu et al., 2019), being heavy (Pu et al., 2019; Robinson et al., 2013), potential risk of infection due to difficulty maintaining hygiene, and the training and time burden for staff (Bemelmans et al., 2013; Moyle et al., 2016). Similar limitations were reported in lower quality studies (Bemelmans et al., 2016, 2013; Robinson et al., 2013) (Appendix 5).

Summary: PARO is acceptable for use in LTC settings and people with dementia generally appear to enjoy using PARO. However, it seems to be less used in a private home setting based on one moderate quality study. Reported perceptions of PARO as childish, noisy, and heavy can limit its acceptability. Infection risk and potential training and time burden on staff can be an obstacle for more widespread adoption in LTC settings.
### Table 1

**Summary of robots.**

<table>
<thead>
<tr>
<th>Category</th>
<th>Name</th>
<th>Appearance</th>
<th>Features</th>
<th>Summary</th>
<th>Estimated cost:</th>
<th>Cost:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pet companion robots</td>
<td>PARO</td>
<td>Appearance: PARO is a battery-charged seal robot covered with white artificial fur. Weight 2.7 kg.</td>
<td>Features: PARO has tactile, light, audio, temperature, and posture built-in sensors. It can interact with the users by making sounds and moving its neck, flippers, and tail. But PARO cannot move forward. <strong>Estimated cost:</strong> US$6675</td>
<td><strong>Humanoid companion robots</strong></td>
<td>Hybrid-face robots</td>
<td>Appearance: Hybrid-Face robot is a teleoperated humanoid robot. It consists of a tablet (with Hybrid-Face software) and a 3D faceplate mounted on a stand. <strong>Features:</strong> It can be controlled remotely using a laptop with internet connection. It can express emotions (anger, sad, afraid, happy, surprised, afraid, disgust) using facial features (raising eyebrows). <strong>Estimated cost:</strong> N/A</td>
</tr>
<tr>
<td></td>
<td>AIBO</td>
<td>Appearance: AIBO is a battery-charged robotic dog which weighs 1.5 kg. It is made of metal. <strong>Features:</strong> AIBO can interact with the user by shaking its body, stretching, raising its paw and barking. AIBO also has the artificial intelligence capability which allows it to develop its own ‘personality’ through interactions with users. It also has a built-in camera to support photo-taking function. <strong>Estimated cost:</strong> US$1876</td>
<td></td>
<td><strong>JustoCat</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JustoCat</td>
<td>Appearance: JustoCat is a battery-charged robotic cat. It has the same size and weight as a real cat and has washable fur. <strong>Features:</strong> JustoCat can interact with the user by shaking its body, raising its paw and a purring. <strong>Estimated cost:</strong> US$1426</td>
<td></td>
<td><strong>NECRO</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NECRO</td>
<td>Appearance: NECRO is a battery charged robotic cat covered with grey fur. Weight 1.6 kg. <strong>Features:</strong> NECRO can interact by making verbal responses (e.g., meow) and nonverbal responses (e.g. moving its paws, wagging tail). NECRO can respond to its name when called. Its artificial intelligence capability also allows it to develop its own personality. However, NECRO cannot walk. <strong>Estimated cost:</strong> N/A</td>
<td></td>
<td><strong>Pleo</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pleo</td>
<td>Appearance: Pleo is a battery charged robotic dinosaur which weighs 1.6 kg. <strong>Features:</strong> It has multiple sensors to see, sense, touch, and detect objects, sense motion and track a moving object and move autonomously. Users can interact with Pleo using physical, verbal and tactile communication. <strong>Estimated cost:</strong> US$518</td>
<td></td>
<td><strong>Hasbro Joy for all</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hasbro Joy for all</td>
<td>Appearance: Hasbro Joy for all companion robots are battery-charged robotic dogs/cats mimicking the appearance of a real dog and cat. <strong>Features:</strong> Their built-in sensors respond to motion and touch. Users can interact with them by petting and hugging. They can move, nuzzle, meow/bark and roll over. <strong>Estimated cost:</strong> US$125</td>
<td></td>
<td><strong>FurReal Friends robot (Daisy Plays with Me Kitty)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FurReal Friends robot (Daisy Plays with Me Kitty)</td>
<td>Appearance: Daisy Plays with Me Kitty is a robotic cat which weighs approximately 0.45 kg. <strong>Features:</strong> It has tactile sensor and movement sensor in her eyes and torso. Users can pet her forehead and body and it will respond to the user by moving, jumping, purring, or meowing. <strong>Estimated cost:</strong> US$4875</td>
<td></td>
<td><strong>Homecare assistive robot</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Telepresence communication robot</strong></td>
<td><strong>Giraff</strong></td>
<td><strong>Vgo</strong></td>
<td><strong>Ed</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Appearance:</strong> Giraff is a wheel-based robot of human height with a large video LCD screen display, video camera, speaker and microphone. Weights 14 kg. <strong>Features:</strong> Giraff can be teleoperated by a caregiver moving the Giraff to make social gestures to interact with someone with dementia (e.g., moving the video display up and down nodding). It needs to be used in indoor environment with a WiFi signal. Estimated cost: US$11,900</td>
<td></td>
<td><strong>Appearance:</strong> Vgo is a wheel-based robot, weighs 9 kg and 122 cm tall. Vgo has a more streamlined appearance and a smaller screen display than Giraff. <strong>Features:</strong> With Vgo, the remote user can hear, talk, interact and go anywhere within the home with the Person with dementia. It needs to be used in indoor environment with a WiFi signal. Estimated cost: US$4875</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued on next page)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Category</th>
<th>Name</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guide</td>
<td>Sophie and Jack/Papero</td>
<td>Appearance: The robot is a humanoid robot with a touch screen on the torso. Can interact with users by using the touch screen and voice interface. Remind users of daily schedule and 5.) making Skype/phone call functions. Estimated cost: N/A</td>
</tr>
<tr>
<td>Multifunctional robot</td>
<td>MARIO</td>
<td>Appearance: MARIO’s appearance loosely mimics a human with a head and a body which has a tablet personal computer. It is capable of multimodal interactions with users. Pepper is recognizes faces and human emotions and can interact via touchscreen and speech commands. Reminding about daily schedule. Estimated cost: N/A</td>
</tr>
<tr>
<td></td>
<td>NAO</td>
<td>Appearance: NAO is a humanoid robot. It is 58 cm tall and weighs 4.3 kg. Features: It contains the ‘search and approach’ function, monitoring function, message notification and phone-call functions. It can be used to remind people with dementia to take medication and provide reminders to caregivers if the people with dementia leaves the house. Estimated cost: N/A</td>
</tr>
<tr>
<td></td>
<td>3DX robot</td>
<td>Appearance: 3DX robot comprises of a gaming laptop (as a head) and a wheel-based stand (torso). Features: It is capable of navigation, communication, and recognizes faces and speech. It contains the ‘search and approach’ function, monitoring function, message notification and phone-call functions. Estimated cost: N/A</td>
</tr>
<tr>
<td></td>
<td>Silbot</td>
<td>Appearance: Silbot is a wheel-based robot weighing approximately 25 kg and is 114 cm tall. Features: Silbot can interact with the user by using the touch screen and voice interface. Estimated cost: N/A</td>
</tr>
<tr>
<td></td>
<td>Ourpuppet ELISA</td>
<td>Appearance: Ourpuppet ELISA is a robotic companion doll with a puppet look. Function: It can monitor the emotional state and health condition of the people with dementia. If there are any unwanted situation (e.g. the people with dementia are agitated), the robot will calm the people with dementia. Estimated cost: N/A</td>
</tr>
</tbody>
</table>

3.3.1.2. Meta-analysis of PARO’s effects on dementia outcomes. The summarised SMD are in Fig. 2 and Appendix 6a-g shows all individual forest plots and meta-analysis. There were no significant effects on any outcome, as detailed below.

3.3.1.2.1. Cognition. Six studies of PARO examined its effect on cognition, three from Spain (Valenti Soler et al., 2015), two from Korea (Koh and Kang, 2018; Song, 2009) and one from New Zealand (Liang et al., 2017). Among these, four were RCTs (Koh and Kang, 2018; Liang et al., 2017; Valenti Soler et al., 2015) and two were non-randomized interventional studies (Song, 2009; Valenti Soler et al., 2015). These studies included 339 people with dementia (85% females and 85% from LTC).

The only moderate quality RCT (Liang et al., 2017) found no immediate or long-term effects of a PARO group intervention on cognition, compared to usual care. The remaining lower or moderate quality RCTs and non-RCTs also showed no effect on cognition when comparing PARO sessions with either a usual care group or a real dog group, in both LTC and community settings (Koh and Kang, 2018; Song, 2009; Valenti Soler et al., 2015). Meta-analysis of the immediate cognitive outcomes of PARO compared to usual control found no difference (three trials; 128 participants; SMD = 0.03; 95% CI: −0.32 to 0.38, I² = 0%) (see Fig. 2 and Appendix 6a for detailed figure). One large trial (Valenti Soler et al., 2015) comparing PARO to active control found no effect on cognition (516 participants; SMD = 0.07; 95%CI: −0.40 to 0.54).

Summary: There is evidence that PARO does not have an immediate or long-term effectiveness cognition.
Table 2
Studies of PARO with moderate and high study quality (N = 11 studies).

<table>
<thead>
<tr>
<th>Author, year, country</th>
<th>Participants</th>
<th>Study design (sample size)</th>
<th>PARO session</th>
<th>Control</th>
<th>Measures on acceptability/feasibility/effectiveness</th>
<th>Finding on acceptability/feasibility/effectiveness</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones et al., 2018 (Jones et al., 2018) Australia</td>
<td>Nursing home people with dementia 73% females; Mean age (SD): 84 (8); Mean RUDAS 7/30 (7)</td>
<td>Quantitative cross-sectional study (138)</td>
<td>30 individual non-facilitated© session (15 mins); 3 times/week for 10 weeks; led by research assistants</td>
<td>N/A</td>
<td>Agitation; CMAI-SF; assessed before and immediately post-intervention</td>
<td>People with dementia with more agitation had higher level of agitation immediately after the intervention.</td>
<td>High</td>
</tr>
<tr>
<td>Moyle et al., 2019, (Moyle et al., 2019) Australia (Qualitative arm (Moyle et al., 2017b))</td>
<td>Family caregivers of nursing home 75% females; age range: 20–65;</td>
<td>Qualitative study (20)</td>
<td>Same as above</td>
<td>N/A</td>
<td>Interviews after observing-robot interaction</td>
<td>Family had more positive views of PARO than plush toy. More feasible to use than real animals; PARO was useful in improving mood and communication; but childish and costly.</td>
<td>High</td>
</tr>
<tr>
<td>Moyle et al., 2018 (Moyle et al., 2016) Australia (Qualitative arm of (Moyle et al., 2017b))</td>
<td>Staff of nursing home 100% female; age range: 20–75; Mostly nursing staff</td>
<td>Qualitative study (20)</td>
<td>Same as above</td>
<td>N/A</td>
<td>Interviews with staff after observing robot interaction</td>
<td>Staff willing to use PARO; felt it was useful for mood and building a positive environment in the home; However childish; costly; noisy, needed staff training; hygienic issues, less acceptable in severe dementia</td>
<td>High</td>
</tr>
<tr>
<td>Hung et al., 2019 (Hung et al., 2019a) Canada</td>
<td>People with dementia in hospital 60% females -Mostly aged 76 or above -Mostly moderate-severe dementia</td>
<td>Qualitative study (10)</td>
<td>An individual facilitated session© (20–30 mins); delivered 2–4 times/week; Led by family co-researchers.</td>
<td>N/A</td>
<td>Interviews during the intervention</td>
<td>PARO was generally accepted by participants who viewed PARO as a companion; Fun; Useful in social connection, humanizing hospital setting</td>
<td>High</td>
</tr>
<tr>
<td>Pu et al., 2019 (Pu et al., 2019) Australia (Same study as (Pu et al., 2020) but only reported qualitative arm)</td>
<td>People with dementia with chronic pain in nursing home 82% females; Mean age(range): 84 (65–94); Mean MMSE (range): 15/30 (9–24)</td>
<td>Qualitative study (11)</td>
<td>30 individual non-facilitated sessions© (15 mins); delivered 5 times/week for 6 weeks</td>
<td>N/A</td>
<td>Interviewed the PWD after the intervention</td>
<td>PWD generally liked about the appearance and interactive feature and most reflected no dissatisfaction; PARO was useful in improving mood and relieve pain; Limitations: Heavy, Noisy, unable to walk</td>
<td>High</td>
</tr>
<tr>
<td>Moyle et al., 2017 (Moyle et al., 2017b) Australia</td>
<td>People with dementia in nursing home average; Mean age (SD): 85 (8); RUDAS mean (SD) = 7/30 (7)</td>
<td>Three-arm cluster-randomized RCT (415)</td>
<td>30 individual non-facilitated© session (15 mins); 3 times/week for 10 weeks; led by research assistants</td>
<td>Plush toy Usual care</td>
<td>1. Attrition rate 2. Average number of attended sessions 3. Agitation; CMAI-SF; assessed before, immediately after and 5 weeks follow up.</td>
<td>1. Attrition rate: 3%© 2. Participants s attended 26/30 sessions on average 3. PARO group were more visually and verbally engaged than the plush toy group and had a greater pleasure and reduced neutral affect than usual care. PARO had less agitation behaviors than the usual care during the session.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Pu et al. (2020) (Pu et al., 2020) Australia</td>
<td>People with Dementia with chronic pain in nursing home 76% females; Mean age (SD): 86 (7); All moderate-severe stage</td>
<td>Two-arm individual randomized RCT (43)</td>
<td>30 sessions of individual non-facilitated intervention© (15 mins) delivered 5 times/week for 6 weeks</td>
<td>Usual care group</td>
<td>1. Attrition rate of PARO sessions 2. Agitation (CMAI-SF); Depression (CSDD); Anxiety(RAID); assessed before and immediately post-intervention</td>
<td>1. Attrition rate: 0%. 2. No immediate effect on agitation, depression and anxiety than usual care group</td>
<td>Moderate</td>
</tr>
<tr>
<td>Joranson et al. (2015a) (Joranson et al., 2015b) Norway</td>
<td>People with dementia in nursing home 67% Females; Mean age (SD): 84 (7); 93% were moderate-severe stage</td>
<td>Two-arm cluster-randomized RCT (60)</td>
<td>24 sessions of group facilitated intervention© (30 mins) delivered twice/week for 12 weeks; led by nurses.</td>
<td>Usual care group</td>
<td>1. Agitation (BABS); Depression (CSDD); assessed before; immediately post-intervention and 3-month follow up</td>
<td>1. No immediate effect, but long-term effect on agitation and depression compared to usual care</td>
<td>Moderate</td>
</tr>
<tr>
<td>Joranson et al. (2016c) (Joranson et al., 2016c)</td>
<td>Same as above</td>
<td>Same as above</td>
<td>Same as above</td>
<td>Same as above</td>
<td>1. Attrition rate of the PARO session© 2. Adherence rate</td>
<td>1. Attrition rate: 10%; Reasons of attrition:</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

(continued on next page)
Table 2 (continued)

<table>
<thead>
<tr>
<th>Author, year, country</th>
<th>Participants</th>
<th>Study design (sample size)</th>
<th>PARO session</th>
<th>Control</th>
<th>Measures on acceptability/feasibility/effectiveness</th>
<th>Finding on acceptability/feasibility/effectiveness</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liang et al., 2017-2017a</td>
<td>People with dementia from New Zealand</td>
<td>Two-arms individual randomized RCT, with mixed method study design (30 People with dementia &amp; 30 family caregivers)</td>
<td>12-18 group facilitated sessions were first carried out in the centre for 6 weeks (two to three times/week; led by researchers), and then a 6-week home-delivered individual PARO 1 intervention was delivered by family CG</td>
<td>Usual care group</td>
<td>1. Attrition rate of the PARO sessions (both day care sessions and home sessions)(^a)</td>
<td>Immediate beneficial effect compared to usual care.</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

Song et al. 2009b | PWD in nursing home – 80% females-Mean age (SD): 85 (8) | Between-subjects pre post study design (40 PWDs) | 12 group facilitated PARO sessions (60 mins); delivered 2 times/week over 6 weeks; Led by researchers | Usual care group | 1. Adherence rate | 1.85% PWD attended 8 out of 12 sessions. | Moderate |

3.3.1.2.2. Neuropsychiatric symptoms. Five studies examined the immediate effect of PARO on overall neuropsychiatric symptoms, three from Spain (Valenti Soler et al., 2015), and one each from New Zealand (Liang et al., 2017) and France (de Sant’ Anna et al., 2012). Among these, three were RCTs (Liang et al., 2017; Valenti Soler et al., 2015) and two were pre/post-test study. (de Sant’ Anna et al., 2012; Valenti Soler et al., 2015) 266 people with dementia were included (84% females and 81% from LTC).

All RCTs (one moderate and two lower quality) found no immediate effects on neuropsychiatric symptoms when compared with usual control (Liang et al., 2017; Valenti Soler et al., 2015) or real dog group. (Valenti Soler et al., 2015) The two lower quality pre/post-test studies, found that PARO had an immediate beneficial effect (de Sant’ Anna et al., 2012) or no effect on neuropsychiatric symptoms (Valenti Soler et al., 2015). Meta-analysis of the immediate neuropsychiatric symptoms outcomes of PARO found no difference compared with usual care (Three trials, 169 participants, SMD = −0.01; 95% CI: −0.32 to 0.29, I\(^2\) =0%) and with active control (two trials; 145 participants; SMD = 0.02; 95% CI: −0.71 to 0.85, I\(^2\) =79%).

**Summary:** There is evidence that PARO does not have an immediate effect on overall neuropsychiatric symptoms. No studies assessed long-term effect.

**Agitation** Five RCTs examined the effectiveness of PARO on agitation. They were from Australia (Moyle et al., 2017b) (Pu et al., 2020), New Zealand (Liang et al., 2017), Norway (Joranson et al., 2015b) and Korea (Koh and Kang, 2018) and included 153 people with dementia. Of these five RCTs, four were of moderate quality (Joranson et al., 2015b; Liang et al., 2017; Moyle et al., 2017b; Pu et al., 2020) and one was of lower quality (Koh and Kang, 2018). One study (Moyle et al., 2017b) was excluded in the meta-analysis as data presented in this study is not sufficient for us to perform meta-analyses as that 10-week follow-up data on the CMAI (the only validated outcome measure which the study used) was missing on a very large and unspecified number of participants.

All the three moderate quality RCTs (one delivered individually and two in groups) found no immediate effect on agitation compared to usual care. (Joranson et al., 2015b; Liang et al., 2017; Pu et al., 2020) However, in the remaining lower quality RCT of PARO delivered in a group setting, an immediate effect was found when compared PARO to the usual care. (Koh and Kang, 2018) Of the two moderate quality RCTs which assessed long-term effect, only one found evidence in improving agitation 12 weeks after the intervention. (Joranson et al., 2015b) Meta-analysis of PARO group compared to usual care found no evidence for effect on agitation immediately after the intervention (4 trials, 153 participants; SMD: −0.25; 95% CI: −0.57 to 0.06, I\(^2\) =0%), or in the long-term (2 trials, 77 participants, SMD = −0.24, 95% CI: −0.94, 0.46, I\(^2\) =0%).
<table>
<thead>
<tr>
<th>Author, year</th>
<th>Study country</th>
<th>Sample Characteristics</th>
<th>Study design</th>
<th>Number</th>
<th>Pet robot intervention</th>
<th>Control group/condition</th>
<th>Measures</th>
<th>Finding on acceptability and feasibility</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gustafsson et al., 2015</td>
<td>Sweden</td>
<td>Care home staff and family caregivers of people with severe dementia and neuropsychiatric symptoms in nursing home</td>
<td>Qualitative study (Part of a mixed method study)</td>
<td>3 family caregivers; 11 staff</td>
<td>A 7-week group facilitated session using Justocat by care staff</td>
<td>N/A</td>
<td>Interview findings from CG and staff</td>
<td>Participants liked appearance and interactive feature; useful; easy to use; feasible to use. But some PWD showed anxiety</td>
<td>Moderate</td>
</tr>
<tr>
<td>Barakova et al., 2020</td>
<td>Netherlands</td>
<td>People with dementia in LTC facility Mean age: 85 (4.8) 75% females, MMSE= 19: 75%</td>
<td>Observational study</td>
<td>16</td>
<td>Participants interacted with a robotic sheep, and then a virtual sheep (shown on the video screen-augmented reality display) and robotic sheep would respond to the users' stroke/touch. Four 20-min individual facilitated weekly session, led by a trained facilitator</td>
<td>-Participants interacted with a robotic sheep that was turned off, accompanied with the same video -Participants watched the video only</td>
<td>1. Observed engagement; Observation measurement of engagement (OME).2. The engagement of a person with dementia (EPWDS).3. Observed emotional rating scale (OERS). All observed during the session</td>
<td>1. The use of augmented reality display (use of virtual sheep) enhanced participants 'experience of robot interaction, with increased attention and interaction observed. But no improvement on positive affect states.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Feng et al., 2021</td>
<td>Netherlands</td>
<td>Same as above</td>
<td>Same as above</td>
<td>Same as above</td>
<td>Two different interventions were used</td>
<td>Same as above</td>
<td>1. Ethnographic and Laban-Inspired Coding System of Engagement (ELICSE).2. Observation measurement of engagement (OME).3. The engagement of a person with dementia (EPWDS).4. Observed emotional rating scale (OERS). All observed during the session</td>
<td>1. When additional auditory modality was programmed in the virtual sheep (augmented reality display), there are more positive attitude, positive behavioural engagement, and communications with the robots.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Schuurmans et al., 2021</td>
<td>Netherlands</td>
<td>People with dementia in nursing home</td>
<td>Cluster individual three-arm RCT</td>
<td>66</td>
<td>Eight 30-min facilitated weekly group sessions using FurReal Friends robot (Daisy Plays-With-Me Kitty); Led by a facilitator</td>
<td>-Facilitator only - Facilitator and a dog</td>
<td>1. Attrition; 2. Behaviours were observed in three of the sessions (beginning, middle, end)</td>
<td>1. Only 25% participants from the robot group (4 out of 16) completed all the sessions. 2. Participants showed more interactions with the facilitator when robot was present. 3. For the latter sessions, there are less interaction with the robot</td>
<td>Lower</td>
</tr>
<tr>
<td>Libin and Cohen-Mansfield, 2016</td>
<td>US</td>
<td>Nursing home people with dementia 100% females Mean age (range): 90 (83-98) Mean GDS: 5.4/7 (4.7)</td>
<td>Before and after interventional study</td>
<td>9</td>
<td>10-minute individual facilitated session using NeCoRo; led by research assistant</td>
<td>Participants were own control -also attended a plush toy cat session</td>
<td>1. Affect; LMBS; 2. Agitated behaviours; AbMI All assessed before and during the session.</td>
<td>1. Increase in pleasure and interest of robot group only. 1. No change in anger and anxiety for either group. 2. Reduction on agitation for plush toy group only.</td>
<td>Lower</td>
</tr>
<tr>
<td>Hammarlund et al., 2021</td>
<td>US</td>
<td>People with dementia using day care services</td>
<td>Before and after interventional study</td>
<td>5</td>
<td>Individual non-facilitated sessions for 4 weeks using Hasbro robotic dogs/cats (free</td>
<td>N/A</td>
<td>1. Agitation; CMAI 2. Quality of life; DEMQOL</td>
<td>1. Improved quality of life, but not agitation, after the sessions</td>
<td>Lower</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Author, year country</th>
<th>Sample Characteristics</th>
<th>Study design</th>
<th>Number</th>
<th>Pet robot intervention</th>
<th>Control group/ condition</th>
<th>Measures</th>
<th>Finding on acceptability and feasibility</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Marx et al., 2010) US</td>
<td>Nursing home people with dementia 79% females Mean age (range): 87 (61–101) MMSE (SD; Range): 9.1/30 (6.2, 0-21)</td>
<td>Observational interventional study (Part of mixed method study)</td>
<td>56</td>
<td>access to robots with no time restriction) An individual non-facilitated session (15 mins at maximum) using a robotic dog; led by a researcher</td>
<td>Participants were own control by attending a video-watching, colouring, plush toy and real dog sessions</td>
<td>All assessed before and after sessions</td>
<td></td>
<td>1. 36% refused robot session 2. Attitudes towards the robot: neutral to somewhat positive; Average engagement with robotic dog ~ 2 mins, with no difference found across different conditions. 3. 15 verbal responses to robotic dog and puppy video, 5 in colouring task, 3 in plush toy. 23-24 with real dog</td>
</tr>
<tr>
<td>(Kramer et al., 2009) US</td>
<td>Nursing home people with dementia 100% females</td>
<td>Observation interventional study</td>
<td>18</td>
<td>Three AIBO-assisted visits (3 mins) to the PWD weekly for three weeks</td>
<td>Participants were own control - Visitor only - Real dog-assisted visit</td>
<td>Count of conversations, touch, smiles and laughs, and looks</td>
<td>More touch in AIBO and dog visit than visitor-alone session No difference in conversation, smiles and laughs and looks across different visits.</td>
<td>Lower</td>
</tr>
<tr>
<td>(Tamura et al., 2004) Japan</td>
<td>Nursing home people with dementia 92% Females Mean age: 84 Severe dementia</td>
<td>Observation interventional study</td>
<td>13</td>
<td>5 min group occupational therapy session using AIBO</td>
<td>Participants were own control by exposing to a toy dog condition.</td>
<td>Behaviours were observed within the session</td>
<td>Less touch, talk, clap, and looks with AIBO than the toy dog</td>
<td>Lower</td>
</tr>
<tr>
<td>(Fogelson et al., 2021) US</td>
<td>Family &amp; Staff of people with dementia from care homes Not reported</td>
<td>Qualitative study (part of a mixed method study)</td>
<td>Not reported</td>
<td>Six 30-min weekly individual non-facilitated session using Joy for all robotic dog/cat (free access to the robots placed inside their room)</td>
<td>N/A</td>
<td>Interview findings from people with dementia, CG &amp; staff</td>
<td>Robotic dogs/cats provided meaningful &amp; positive experience to the users; Highly engaged; Provided companionship; Facilitated communication.</td>
<td>Lower</td>
</tr>
<tr>
<td>(Tummers et al., 2020) Netherlands</td>
<td>People with dementia from care homes Not reported</td>
<td>Mixed method study - Observational study, and individual interviews with people with dementia</td>
<td>12</td>
<td>Pleo Dinosaur with a facilitator - Interaction with a dog and a facilitator - Facilitator only</td>
<td>1. Observed engagement levels 2. Interviews findings with people with dementia</td>
<td></td>
<td>1. Active engagement and positive interactions observed in Pleo session. 2. Pleo triggered more interactions &amp; conversations than the facilitator-only sessions, but similar to the dog session</td>
<td>Lower</td>
</tr>
</tbody>
</table>

ABMI: Agitated Behaviors mapping instrument; CDR.CMAI: proxy-rated Cohen-Mansfield Agitation Inventory; DEMQOL: Quality of Life Assessment in Dementia; GDS: Global deterioration scale; LMBS: Lawton’s modified behaviour stream; MMSE: Mini-mental status examination; MSQ: Mental status examination; OME: observational measurement of engagement CG: caregiver; LTC: long-term care; PWD: person with dementia; Justocat, NeCoRo and AIBO were names of robots.

* Instruments that had been validated on PWD are highlighted in purple. Other non-highlighted information is either non-validated instruments or trial data (e.g., dropout rate) ** Significant results are highlighted in bold
Table 4  
Studies of humanoid companion robots (n = 4).

<table>
<thead>
<tr>
<th>Study design</th>
<th>Control group</th>
<th>Measures</th>
<th>Finding</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two arm RCT, with an ABAB withdrawal design</td>
<td>Usual care</td>
<td>1. Neuropsychiatric symptoms; NPI-Q. 2. Quality of life; QoL-AD</td>
<td>1. No improvement on neuropsychiatric symptoms and quality of life compared to the usual care group</td>
<td>Moderate</td>
</tr>
<tr>
<td>Focus group interviews (PwD N = 1, staff n = 2, CG–4, robotic experts n – 2) &amp; individual interviews with PWD-caregiver dyads (N = 14)</td>
<td>Nil</td>
<td>Interview findings</td>
<td>- Easy to operate; facilitates engagement; appearance and voice of are acceptable. -Limitations: High cost, cannot be used on non-English speaker, potentially reduced human contact, health monitoring feature is preferred</td>
<td>Moderate</td>
</tr>
<tr>
<td>Observational study (19 PwDs)</td>
<td>A teleoperated Lugwid robot conversed with PWD about a picture</td>
<td>1. Usability questionnaire; assessed before and after the robot session 2. Audio recording and facial tracking data of PWD collected in the sessions</td>
<td>1. Compared to facilitator, there is a shorter conversation and less engagement and more misunderstanding when PWD interacting with the robots; 2. Mixed opinions regarding the likeliness and eagerness to interact with the robot 3. PWD showed more difficulty in understanding the autonomous robot compared to the teleoperated one 4. Limitations: robotic voice, lack of prosodic, gestural, and facial cues of the robots</td>
<td>Lower</td>
</tr>
</tbody>
</table>

PwD: People with dementia  
* Significant results highlighted in bold  
* Instruments which are under the section of ‘Measures’ hat had been validated on PwD are highlighted in purple. Other non-highlighted information is either non-validated instruments or trial data (e.g., dropout rate).

=53% (see Fig. 2, Appendix 6c for details).

Summary: There is evidence that PARO does not have an immediate (neither individual nor group) or long-term effect in reducing agitation.

Anxiety: Three RCTs examined PARO’s immediate effects on anxiety, (Moyle et al., 2013a; Petersen et al., 2017; Pu et al., 2020) two with usual care and one compared to a reading group. Two were from Australia (Moyle et al., 2013a; Pu et al., 2020) and one from US (Petersen et al., 2017). 122 people with dementia participated (Mean age: 85, 100% from LTC). The moderate quality RCT (Pu et al., 2020) found individual PARO intervention had no effect on reducing anxiety compared to usual care. The remaining two lower quality RCTs, however, found an immediate effect on reducing anxiety when comparing PARO with a usual care group and a reading group (Moyle et al., 2013a; Petersen et al., 2017). Meta-analysis of the two trials compared with usual care with a total of 104 participants found no evidence of an effect compared to usual care (SMD = 0.24; 95% CI: –0.85 to 1.33; I² = 86%) (see Fig. 2, Appendix 6d for details).

Summary: There is evidence that PARO does not have an immediate effect on anxiety but no high-quality studies. No studies examined long-term effect.

Depression: Five studies examined PARO’s effect on depression, one each from Australia, (Pu et al., 2020) New Zealand, (Li et al., 2017) Norway, (Joranson et al., 2015b) US (Petersen et al., 2017) and France (de Sant’Anna et al., 2012). Among these, four were RCTs (Joranson et al., 2015b; Li et al., 2017; Petersen et al., 2017; Pu et al., 2020) and one was a pre/post-test study. (de Sant’Anna et al., 2012) 199 people with dementia participated (Mean age: 84; 71% females; 85% from LTC). The three moderate quality RCTs (Joranson et al., 2015b; Li et al., 2017; Pu et al., 2020) reported no immediate effect on depression. The two RCTs which included a longer follow-up timepoint (Joranson et al., 2015b; Li et al., 2017) had differing findings, with one showing a significant long-term effect on depression 12 weeks after the intervention, the other which assessed long term effect 6 weeks after the intervention did not. Results from the remaining two lower quality studies were mixed. A non-blinded RCT, (Petersen et al., 2017) but not the pre/post-test study, (de Sant’Anna et al., 2012) found PARO was effective in reducing depression compared to the usual care group immediately after the intervention. Meta-analysis found no immediate effect.
effect (4 trials, 181 participants, SMD = 0.08, 95% CI: -0.52 to 0.69, $I^2 = 75\%$) or long-term effect (2 trials, 77 participants, SMD = 0.01, 95% CI = -0.75 to 0.77, $I^2 = 59\%$) on depression (see Fig. 2, Appendix 6e (i-ii) for details).

**Summary** There is evidence that PARO does not have an immediate or long effect on depression but no high-quality studies.

**Apathy** Five lower quality studies examined the immediate effect on apathy. Three were from Spain (Valenti Soler et al., 2015) and one each from Australia (Moyle et al., 2013a) and France (de Sant’Anna et al., 2012). Among these, three were RCTs (Moyle et al., 2013a; Valenti Soler et al., 2015) and two were non-RCTs (de Sant’Anna et al., 2012; Valenti Soler et al., 2015). 251 people with dementia participated (87% females: 86% from LTC facilities).

The three RCTs reported mixed results. One study (Valenti Soler et al., 2015) showed immediate beneficial effect on apathy when comparing PARO to usual care, but not when compared to a multi-functional robot. No effect on apathy was found in the other two RCTs which compared PARO to a reading group (Moyle et al., 2013a), a real dog group and a usual care group (Valenti Soler et al., 2015). The two pre/post-test studies found no change in apathy (de Sant’Anna et al., 2012; Valenti Soler et al., 2015). Meta-analysis of two lower quality RCTs which compared PARO to active control (reading group and multifunctional robot) did not find an effect on apathy (81 participants, SMD = 0.14, 95% CI: -0.29 to 0.58, $I^2 = 0\%$) (See Fig. 2, Appendix 6f for details).

**Summary** There is evidence that PARO does not have an immediate or long-term effect on apathy but no high-quality studies, and no studies assessed long-term effects.

**Quality of life** Four RCTs (one moderate quality and three lower quality) examined PARO’s effect on quality of life (Jøranson et al., 2016b; Moyle et al., 2013a; Valenti Soler et al., 2015). Two were from Spain (Valenti Soler et al., 2015) and one was from each of Norway (Jøranson et al., 2016b) and Australia (Moyle et al., 2013a). 289 participants participated (84% females, 100% from LTC). The moderate-quality RCT (Jøranson et al., 2016b) reported no immediate or long-term effects on quality of life for PARO when compared with a usual care group. Two of three remaining lower quality RCTs reported beneficial effects on quality of life when comparing PARO to usual care, real dog (Valenti Soler et al., 2015) and reading groups. (Moyle et al., 2013a) Meta-analysis (Fig. 2, Appendix 6g (i-ii) for details) of an immediate effect on quality of life compared to usual care (2 trials, 127 participants SMD = -0.05; 95% CI: -0.52 to 0.42, $I^2 = 43\%$) or a real dog intervention or reading group (2 trials, 159 participants, SMD = -0.26, 95%CI: -0.76 to 0.05, 12 =0%) found no evidence of a difference although numbers were small.

**Summary** There is evidence that PARO does not have an immediate
on quality of life but no high-quality studies. Only one moderate quality study assessed long-term effects on quality of life and produced no evidence of positive effects.

3.3.2. Other companion robots robotic animals

11 papers examined robotic animals (Table 3) (two papers reported the same study). One was RCT from Netherlands (Schuurmans et al., 2021), seven were interventional/observational studies (Barakova et al., 2020; Feng et al., 2021; Hammarlund et al., 2021; Kramer et al., 2009; Libin and Cohen-Mansfield, 2016; Marx et al., 2010; Tamura et al., 2004) from US, Netherlands, and Japan, and one a qualitative study from Sweden and US (Fogelson et al., 2021; Gustafsson et al., 2015), one a mixed method study from Netherlands (Tummers et al., 2020). 10 assessed feasibility and acceptability and one assessed effectiveness. A total of 195 people with dementia (97% from LTC), three family caregivers, and 11 healthcare professionals were included. Caregivers and staff were interviewed in one moderate quality qualitative study (Gustafsson et al., 2015). Justocat, a robotic cat, was found to be highly acceptable due to its appearance, interactive features, and usefulness. However, some participants with dementia were anxious in its presence. Two moderate quality papers reporting the same observational study found that people with dementia’s user experience could be enhanced when augmented virtual reality function was used (a virtual sheep shown on screen would respond to users’ stroke, touch and speech towards the robotic sheep) (Barakova et al., 2020; Feng et al., 2021), leading to more observed attention and interaction.

The remaining seven lower quality studies reported conflicting findings in observed engagement and affect (Fogelson et al., 2021; Kramer et al., 2009; Libin and Cohen-Mansfield, 2016; Marx et al., 2010; Schuurmans et al., 2021; Tamura et al., 2004; Tummers et al., 2020) Four studies found there were more interactions or engagements during the robotic animals session, compared to the facilitator-only session (Kramer et al., 2009; Schuurmans et al., 2021; Tummers et al., 2020) and plush toy/colouring sessions (Marx et al., 2010). But when compared to a session with a real-dog, two studies (Tummers et al., 2020) (Marx et al., 2010) found no difference on interactions/engagement level. Another study found that a NeCoRo, a robotic cat, led to more positive affect but no improvement on agitation behaviors compared to control conditions. (Libin and Cohen-Mansfield, 2016) Two studies reported high dropout: one (Schuurmans et al., 2021) found 75% participants failed to complete all the eight robot sessions, and another one (Marx et al., 2010) found 36% of participants refused to attend the robotic dog session, but those who attended had a neutral to positive attitude towards the robotic dog. One study found that people with severe dementia interacted less with the robotic dog than a toy dog (Tamura et al., 2004). The only lower quality study reported positive findings in general, suggesting robotic dogs/cats provided meaningful and engaging companionship experience to people with dementia in LTC setting (Fogelson et al., 2021).

Only one lower quality before and after interventional study (Hammarlund et al., 2021) examined effectiveness. Robotics dogs or cats were found effective in improving quality of life, but not agitation in people with dementia.

Summary A single moderate-quality qualitative study showed that family caregivers and staff were positive about the use of a robotic cat in a nursing home, with two moderate quality studies suggested the use of augmented virtual reality function can further enhance the user experience. The evidence was unclear on engagement with robotic animals, with the lower quality studies reporting mixed results. The evidence of effectiveness is unclear as there was only one lower quality study which reported improvement on quality of life but not agitation after sessions.

3.3.3. Humanoid companion robots

Four papers (two papers reporting the same study) examined humanoid companion robots (Table 4). Two RCTs reported the same study from Hong Kong (Chen et al., 2020; Ke et al., 2020): one was a qualitative study from India (Natarajan et al., 2021) and the other an observational study from Canada (Pou-Prom et al., 2020). There were two moderate quality (Chen et al., 2020; Natarajan et al., 2021) and two lower quality studies (Ke et al., 2020; Pou-Prom et al., 2020). Three studies reported feasibility/acceptability data (Ke et al., 2020; Natarajan et al., 2021; Pou-Prom et al., 2020) and one reported effectiveness (Chen et al., 2020). They included a total of 130 people with dementia (94% from LTC), 11 family caregivers, 2 dementia staff and 2 robotic experts.

Three studies examined feasibility and acceptability of humanoid companion robots. Overall, humanoid companion robots were acceptable. In the moderate quality qualitative study (Natarajan et al., 2021), people with dementia, family caregivers, staff, and robotic experts liked the humanoid appearance and human voice of the robots, and perceive it as easy to operate and useful to facilitate engagement. However, they reported limitations such as high cost, language restriction (only English), lack of homecare functions and reduced human contacts. For the lower quality studies, one observational study found that using a robot was associated with better technological skills (Ke et al., 2020) and, in another study of a humanoid companion robot used to discuss a picture with people with dementia, (Pou-Prom et al., 2020) there was less engagement and more misunderstanding in the robots session (Lugwid) than the facilitator-only session.

Only one moderate quality RCT examined the effectiveness of humanoid companion robot (Chen et al., 2020). Kabochan was not more effective than usual care in improving neuropsychiatric symptoms and quality of life.

Summary Humanoid companion robots appeared feasible and acceptable to people with dementia, caregivers, and staff, but cost and potential reduction of human contact were raised as concerns. When they were used for more advanced tasks, the feasibility was lower. Evidence of effectiveness is solely based on one moderate quality RCT, which found no evidence of benefits of neuropsychiatric symptoms and quality of life.

3.4. Telepresence communication robots

There were three moderate or lower quality studies examining feasibility and acceptability (Table 5). All were from Australia using mixed methods or qualitative interviews, (Moyle et al., 2014, 2018, 2020) two from a single study. (Moyle et al., 2014, 2020) They included ten people with mild to moderate dementia (mean age 82, 40% females, 50% from LTC), 11 family caregivers (mean age 62, 82% females,) and 19 healthcare professionals (mean age 42, 100% females). All studies asked the family caregivers to make a video call to the people with dementia with the robot, and the amount of interaction during these calls varied.

Overall, telepresence robots were acceptable. In the moderate quality qualitative study, family caregivers perceived telepresence robots to be useful and appreciated their mobility. (Moyle et al., 2020) Video observational data from lower quality mixed method studies (Moyle et al., 2014, 2018) also found participants with dementia showed a positive affect and were engaged in 93% of calls. They reported having an authentic experience when using the telepresence robots, although the instrument used to assess this had not been validated with people with dementia. (Moyle et al., 2018) Despite high acceptability, there were frequent barriers including technological challenges, especially related to internet connectivity, (Moyle et al., 2014, 2018, 2020) difficulty in set-up/operation, (Moyle et al., 2014, 2018) privacy concerns, (Moyle et al., 2018) time commitment for family caregivers and staff, and being impractical for severe dementia. (Moyle et al., 2014, 2020).

Summary Telepresence communication robots were generally acceptable for mild to moderate dementia, but technical challenges were barriers affecting feasibility of use. No studies examined effectiveness of telepresence communication robots.
Three studies were from a single study of a particular robot, Ed, in Canada and took place in a laboratory. (Begum et al., 2015; Rudzicz et al., 2015; Wang et al., 2017) and one was a qualitative study interviewing family caregivers’ perception towards 3DX robot. There were two high quality qualitative studies and two were lower quality.

### 3.5. Homecare assistive robots

We identified four papers about homecare assistive robots (Table 6). Three studies were from a single study of a particular robot, Ed, in Australia. The fourth study took place in Canada and took place in a laboratory. (Begum et al., 2015; Rudzicz et al., 2015; Wang et al., 2017) and one was a qualitative study interviewing family caregivers’ perception towards 3DX robot. There were two high quality qualitative studies and two were lower quality.
<table>
<thead>
<tr>
<th>Author, date, country</th>
<th>Participants</th>
<th>Setting</th>
<th>Study design</th>
<th>Sample size</th>
<th>Homecare robot condition/group</th>
<th>Measures a</th>
<th>Finding</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wang et al., 2017 (Wang et al., 2017) Canada</td>
<td>Mild to moderate PWD 60% females, mean age (SD): 78 (10) Mean MMSE (SD): 21/30 (5) Family CG, Mean age (SD): 60 (13) Either offspring/spouse</td>
<td>Simulated home environment in a rehabilitation Institute;</td>
<td>Qualitative study</td>
<td>10 PWD; 10 family CG</td>
<td>PWD was assisted by Ed to complete a handwashing and tea-making task (12 mins)</td>
<td>N/A Interview findings with PWD and CG</td>
<td>Robot was generally accepted by CG than to PWD. Most PWD did not want a robot due to a lack of perceived need but they are opened to the idea; Limitation: unable to work effectively for complex activity (administering eye-drops), questionable cost-effectiveness, concerns in reducing interaction with CG.</td>
<td>High</td>
</tr>
<tr>
<td>(Arhanat et al., 2020) US</td>
<td>Family caregivers of people with dementia 75% females, mean age (range): 79 (59–98), 50% spouse caregivers</td>
<td>Not reported</td>
<td>Qualitative interviews</td>
<td>8 family CG</td>
<td>3DX robot; Robot performed 2 care protocols related to PWD’s daily routine &amp; home safety (3 h). Caregivers were asked to observe.</td>
<td>N/A Interview findings with CG</td>
<td>Robot was generally accepted by CG. Robot has the potential to reduce caregiving burden by engaging PWDs. Limitations: Navigability in tight space at home, high cost, potential technical problems, robotic voice and non-humanoid features, fear of lack of knowledge towards technology</td>
<td>High</td>
</tr>
<tr>
<td>Begum et al. (2015) (Begum et al., 2015) Canada (This study is the quantitative arm of (Wang et al., 2017))</td>
<td>Same as above for PWD’s sample characteristics</td>
<td>Same as above</td>
<td>Observation intervention study</td>
<td>Same as above for PWD sample size</td>
<td>Same as above</td>
<td>N/A</td>
<td>1. Interaction with the robot 2. PWD’s team behaviors with the robots 3. Positive and negative emotions 4. PWD’s attention and memory All were observed during the session</td>
<td>Lower</td>
</tr>
<tr>
<td>Rudzicz et al., 2015 (Rudzicz et al., 2015) Canada (This study is the quantitative arm of (Wang et al., 2017))</td>
<td>Same as above for PWD’s sample characteristics</td>
<td>Same as above</td>
<td>Same as above</td>
<td>Same as above for PWD sample size</td>
<td>Participants served as their own control by participating in a conversation with an interviewer.</td>
<td>1. PWD’s TIBs were observed during the session and compared with the control condition 2. Speech recognition</td>
<td>1. There was less TIBs in the robot condition than in the human condition, except that PWD expressed more request of non-specific repetition and 'lack of uptake'/ (continued on next page)</td>
<td>Lower</td>
</tr>
</tbody>
</table>
observational studies. They included ten people with mild-moderate dementia (mean age 78, 68% females) and 18 caregivers (mean age:70).

The higher quality qualitative study reported that caregivers generally found Ed more acceptable than did people with dementia (Vang et al., 2017). Most people with dementia perceived they do not need to be assisted but caregivers showed higher willingness to use Ed due to its usefulness. There were concerns about its cost-effectiveness and applicability in complex activities. The two lower quality observational studies consistently reported that although people with dementia showed some positive interactions, there were more frequent negative reactions such as anxiety towards Ed and they were often unwilling to use it. (Begum et al., 2015; Rudzicz et al., 2015).

The high-quality qualitative study of 3DX robot (Arthanat et al., 2020) reported that caregivers found 3DX robot was useful in engaging people with dementia and so could potentially reduce caregiver burden. But they raised concerns about the cost, navigability of robot, and potential technical problems. They also preferred the robot to have humanoid features and voices.

Summary There is little evidence for homecare assistive robots. Only two types of robot have been studied so far with mixed results on feasibility and acceptability on two robots across the four studies. Caregivers appeared to find homecare assistive robots more acceptable than people with dementia. The high cost, technical challenges, navigability, and non-humanoid appearance of robots were factors affecting its acceptability. No studies examined effectiveness of homecare assistive robots.

3.6. Multifunctional robot

15 studies on multifunctional robots (MARIO, Betty, Matilda, Sophie, Jack, PaPeRO, Nao Guide, Silbot, Pepper, Ourpuppet) were identified (Table 7). 13 of these examined feasibility and acceptability, (Casey et al., 2020; Chu et al., 2016; D’Onofrio et al., 2019; Inoue et al., 2012; Khosla et al., 2019, 2017; Kouroupetroglou et al., 2017; Law et al., 2019; Nakamura et al., 2021; Rico et al., 2020; Robinson et al., 2013; Zuschnegg et al., 2021) and three studied effectiveness (D’Onofrio et al., 2019; Valenti Soler et al., 2015).

3.6.1. Feasibility and acceptability

Of the 13 studies examining feasibility and acceptability, (Chu et al., 2016; D’Onofrio et al., 2019; Inoue et al., 2012; Khosla et al., 2019, 2017; Kouroupetroglou et al., 2017; Law et al., 2019; Nakamura et al., 2021; Rico et al., 2020; Robinson et al., 2013; Schramek et al., 2021; Zuschnegg et al., 2021) three were from Australia, (Chu et al., 2016; Khosla et al., 2019, 2017) two from New Zealand (Law et al., 2019; Robinson et al., 2013), six from Europe (one study in two countries Italy and Ireland, and the other two studies in three countries Italy, Ireland and UK, one study each in Spain, Germany, Austria) (Casey et al., 2020; D’Onofrio et al., 2019; Kouroupetroglou et al., 2017; Rico et al., 2020; Schramek et al., 2021; Zuschnegg et al., 2021), and two from Japan (Inoue et al., 2012; Nakamura et al., 2021). Among these, three were high-quality qualitative study (Casey et al., 2020; Law et al., 2019; Zuschnegg et al., 2021) and ten were lower quality observational, intervention or qualitative studies. (Chu et al., 2016; D’Onofrio et al., 2019; Inoue et al., 2012; Khosla et al., 2019, 2017; Kouroupetroglou et al., 2017; Nakamura et al., 2021; Rico et al., 2020; Robinson et al., 2013; Schramek et al., 2021) Studies included a total of 437 people with dementia (57% females), 59 family caregivers (69% females), 91 staff (69% females) and 35 dementia experts (89% females). Most participants were from LTC settings.

Multifunctional robots were generally acceptable to people with dementia, with positive feedback in the three high-quality qualitative study from dementia experts, caregivers, and staff. Experts reported Silbot was simple to use and useful in assisting people with dementia with daily activities (Law et al., 2019). Similarly, Pepper and MARIO were perceived by caregivers and staff to be useful in facilitating communication and interactions (Casey et al., 2020; Zuschnegg et al., 2021), and assisting daily care activities, such as safety monitoring (Zuschnegg et al., 2021). But concerns that robots might replace real human interactions were raised as limitations in both studies (Casey et al., 2020; Zuschnegg et al., 2021).

Positive findings about acceptability was also reflected in six of ten lower quality observational studies/qualitative study which reported people with dementia were largely engaged, enjoyed the social interaction with the robot, and gave frequent positive comments during the session. (Chu et al., 2016; Inoue et al., 2012; Khosla et al., 2017; Nakamura et al., 2021; Rico et al., 2020; Schramek et al., 2021) Other studies reported low feasibility and acceptability due to low engagement (Khosla et al., 2019), technical difficulties (Rico et al., 2020) (D’Onofrio et al., 2019; Kouroupetroglou et al., 2017), including being more difficult to use than PARO. (Robinson et al., 2013) Other reported limitations of multifunctional robots included speech being difficult to understand, (Kouroupetroglou et al., 2017; Law et al., 2019; Robinson et al., 2013), tablet located on the torso of the robot being difficult to use (Rico et al., 2020), overly frequent prompting (Kouroupetroglou et al., 2017; Law et al., 2019), poor ergonomic qualities (Kouroupetroglou et al., 2017; Law et al., 2019; Robinson et al., 2013), and high cost. (Schramek et al., 2021).

Table 6 continued

<table>
<thead>
<tr>
<th>Author, date, country</th>
<th>Participants</th>
<th>Setting</th>
<th>Study design</th>
<th>Sample size</th>
<th>Homecare robot</th>
<th>Control condition/group</th>
<th>Measures</th>
<th>Finding</th>
<th>Study quality</th>
</tr>
</thead>
</table>

| C. Yu et al. | Aging Research Reviews 78 (2022) 101633 |

MMSE: Mini-mental status examination; TIBs: Trouble-indicating behaviors refers to behaviors which reflect PWD’s communication difficulties throughout a conversation. There are 12 types of TIB. For details please refer to (Rudzicz et al., 2015)

CG: caregiver; ED was the name of the homecare assistive robots; PWD: person with dementia

|| Significant results highlighted in bold
For details of study quality rating for each study, please refer to appendix 4a-c.

* Instruments which are under the section of ‘measures’ hat had been validated on PWD are highlighted in purple. Other non-highlighted information is either non-validated instruments or trial data (e.g., dropout rate)
Table 7
Studies of the use of multi-functional robots (N = 15 studies).

<table>
<thead>
<tr>
<th>Author, year, Country</th>
<th>Participants</th>
<th>Setting</th>
<th>Study design (sample size)</th>
<th>Multi-functional robot</th>
<th>Control</th>
<th>Measures *</th>
<th>Finding</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Law et al., 2019 (Law et al., 2019) New Zealand</td>
<td>Experts in elderly care – 79% females’ Diverse occupation</td>
<td>Retirement facility and university</td>
<td>Qualitative study (19 experts)</td>
<td>Silbot</td>
<td>N/A</td>
<td>Study 1: Interview with experts after they watched the prototype video</td>
<td>Silbot was perceived easy to use; stimulating, interacting, useful. But, loud &amp; rough movements; technical issues, patronizing; fast and complex speech, limited functions (e.g., no music activities); not suitable to different dementia severity</td>
<td>High</td>
</tr>
<tr>
<td>Casey et al., 2020 (Italy, UK &amp; Ireland)</td>
<td>PWD in LTC, hospital &amp; community setting Most are mild-moderate dementia; at least 45% females Family CG -69% females Staff &amp; managers-63% females</td>
<td>A geriatric unit of hospital in Italy; LTC facilities in Ireland; community setting in UK</td>
<td>Qualitative study (38 PWD, 28 family CG, 41 staffs &amp; mangers)</td>
<td>MARIO</td>
<td>N/A</td>
<td>1. No. of interactions with robots &amp; Mean total duration of interactions</td>
<td>1. Average no. of interactions (SD; Range): 5.13 (3.44; 1–12) 2. Mean total duration of interactions (SD; range): 199 mins (101; 151–524) 3. Positive views in general: Liked the user-led design and personalization feature; enhanced social interactions &amp; engagement, supported autonomous choices. Limitation: poor voice recognition; absence of monitoring &amp; assessment device; concern about replacing human interactions.</td>
<td>High</td>
</tr>
<tr>
<td>Zuschnegg et al., 2021 (Austria)</td>
<td>Family CG 94% females; Mean age: 54.3 Nurses: 85% females Mean age: 41.1 Dementia trainers 100% females Mean age: 45.9</td>
<td>Non-profit organizations &amp; LTC facilities</td>
<td>Qualitative study (16 CG, 20 nurses, 16 dementia trainers)</td>
<td>Pepper</td>
<td>N/A</td>
<td>Interview findings</td>
<td>Pepper was acceptable: facilitates communication/contacts with others, provides recreational activities, assist activities of daily living, and avoid dangers. But limitations include losing interactions with human &amp; inability to protect PWDs during emergencies</td>
<td>High</td>
</tr>
<tr>
<td>Soler et al. 2015 (Spain)</td>
<td>Mild to severe PWD 88% females Mean age 85</td>
<td>Nursing home</td>
<td>Three arm individual-randomized RCT (101 PWDs)</td>
<td>PARO</td>
<td>Usual care</td>
<td>1. Cognition; MMSE and s-MMSE 2. Neuropsychiatric symptoms, NPI 3. Apathy; APADEM-NH 4. Quality of life; QUALID All were assessed before and after</td>
<td>-PWD in NAO group, but not PARO and usual care, had worsening effect on cognition (MMSE) -NAO group, but not usual care, showed reduction in apathy - No effect on neuropsychiatric symptoms, quality of life and cognition (s-MMSE). Perceived as a good companion; favourable appearance; intention to use again; especially enjoyed the music and game function. But there was operation difficulty(e.g., fast speech, frequent prompting, touch-screen functions) 1. Active engagement (talks, attentive) and positive attitudes. But acceptability was slightly varied across the three countries. Some experienced operational difficulty: 2.2.3. No immediate effect on depression and quality of life</td>
<td>Lower</td>
</tr>
<tr>
<td>Kouroupetroglou et al. (Ireland and Italy)</td>
<td>PWD &amp;Diverse dementia severity</td>
<td>Nursing home and hospital</td>
<td>Intervenational study with post-trial survey mainly (at least 5 PWDs)</td>
<td>MARIO</td>
<td>PWD’s acceptability towards MARIO; surveys completed by PWD and researchers-observing at post-trial</td>
<td>PWD’s acceptability towards MARIO; surveys completed by PWD and researchers-observing at post-trial</td>
<td>Perceived as a good companion; favourable appearance; intention to use again; especially enjoyed the music and game function. But there was operation difficulty(e.g., fast speech, frequent prompting, touch-screen functions) 1. Active engagement (talks, attentive) and positive attitudes. But acceptability was slightly varied across the three countries. Some experienced operational difficulty: 2.2.3. No immediate effect on depression and quality of life</td>
<td>Lower</td>
</tr>
<tr>
<td>D’Onofrio et al., 2019 (Ireland, Italy and UK)</td>
<td>PWD 53% from Italy, 26% from Ireland; 21% from UK63% females Mean age (SD): 77 (10)</td>
<td>Nursing home, hospital and community testing venue</td>
<td>Before-after interventional study (38 PWDs)</td>
<td>MARIO</td>
<td>Quality of Life; QOL-AD; before &amp; after</td>
<td>MARIO’s application was shown to PWD by the researchers and PWD were encouraged to interact with it</td>
<td>1. PWD’s engagement, attention, attitudes, cognitive difficulty, and talks; assessed by OME during the session 2. Depression; GDS; before &amp; after 3. Quality of Life; QOL-AD; before &amp; after</td>
<td>Lower</td>
</tr>
<tr>
<td></td>
<td>A day care centre</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7 (continued)

<table>
<thead>
<tr>
<th>Author, year, Country</th>
<th>Participants</th>
<th>Setting</th>
<th>Study design (sample size)</th>
<th>Multi-functional robot</th>
<th>Control</th>
<th>Measures *</th>
<th>Finding</th>
<th>Study quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>608 Soler et al. 2015 (Valenti Soler et al., 2015) Spain</td>
<td>Community-dwelling PWD 50% females Mean age (range): 78 (68-87)</td>
<td>Before and after interventional study (20 PWDs)</td>
<td>For mild-moderate PWD, 24 group NAO sessions (30-40 mins) were delivered by therapists twice per week over three months. For severe PWD, individual NAO interventions were delivered. Betty was installed at PWD’s house for 3 months.</td>
<td>N/A</td>
<td>1. Cognition, MMSE/2. Neuropsychiatric symptoms, NPI3. Apathy: AI All assessed before and after NAO.</td>
<td>There was a reduction in neuropsychiatric symptom score, but not on cognition (both MMSE and S-MMSE) and apathy after the session with NAO.</td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>Khosla et al., 2019 (Khosla et al., 2019) Australia</td>
<td>PWD Service users of an aged care facility Age range: 75-85</td>
<td>Interventional study, with a post-trial survey mainly (5 PWDs)</td>
<td>N/A</td>
<td>1. Time spent with Betty, 2. Engaged interactions during the session; VC-IOE 3. PWD’s acceptability towards the robot; survey completed by PWDs at post-trial.</td>
<td>1. Each PWD used Betty 5 times/day. 2. Engagement was low: 33.41% of the interactions were behaviourally and visually engaged; 8% for positive engagement; 1% for negative engagement; Most engaged in the singing and dancing application. 3. PWD Enjoyed; found useful; no anxiety. But only 40% wanted to use it again.</td>
<td>Lower</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khosla et al., 2017 (Khosla et al., 2017) Australia</td>
<td>PWD - 70% – 79% females - With different dementia severity</td>
<td>Residential aged care facilities</td>
<td>Observation interventional study (115 PWDs), with a post-trial survey (43 PWDs)</td>
<td>N/A</td>
<td>1. Engagement; VC-IOE; observed at baseline year and then each year for three years 2. PWD’s acceptability towards the robot; survey completed by PWD at post-trial,</td>
<td>There was increase in positive emotional engagement, visual engagement, behavioural engagement throughout the three years, but not on negative emotional engagement and verbal engagement. 2. PWD Enjoyed; found useful, with only 4% PWD showed anxiety towards Matilda. 1. Positive behaviors of PWD increased (Approached to the robot more frequent, more pleasure and positive interaction with robot &amp; others). 2. Staff involvement in PWD activities and staff attention to PWD increased</td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>Chu et al., 2016 (Chu et al., 2016) Australia</td>
<td>PWD: 33% females Aged 65–90 Different dementia severity Care home Staff 64% females</td>
<td>Residential aged care facilities</td>
<td>Observation interventional study (139 PWDs; 25 staff)</td>
<td>N/A</td>
<td>1. PWD’s positive behaviours2. Staff behaviours All were collected at baseline year and then each year for 4 years</td>
<td></td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>Takenobu et al., 2012 (Inoue et al., 2012) Japan</td>
<td>PWD 100% females Mean age (SD; 86 (7) MMSE range: 16–23</td>
<td>A nursing home</td>
<td>Observation interventional study (5 PWDs)</td>
<td>N/A</td>
<td>1. The alert interaction rate of PWD increased 2. Count of positive and negative comments made by PWD within the session 1. Alert interaction rate was 100%; information acquisition rate was 90% 2. 27 positive comments &amp; 4 negative comments. Negative comments: poor ease of use and childishness.</td>
<td></td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>Robinson et al., 2013 (Robinson et al., 2013) New Zealand</td>
<td>PWD 50% females age range:71–93 Family CG 36% females age range: 42–88 Staff 100% females Age range: 45–66</td>
<td>A Dementia residential facility</td>
<td>Observation interventional study (10 PWDs), with qualitative interviews (11 family CG, 5 staff)</td>
<td>PWDs interacted with GUIDE for 5–10 min; led by researchers PWDs interacted with PARO in the same format as GUIDE</td>
<td>1. Behaviors of PWD; observed during the session 2. Interviews with staff and family</td>
<td>1. PWD smiled, touched, and talked less to GUIDE than PARO. 2. Valued ‘multifunctional’. But considered not suitable for PWD due to its poor ease of use (Small image, fast speech) and ergonomics (too tall) They liked the music application the most, but least for the blood pressure function.</td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>(Rico et al., 2020) Spain</td>
<td>PWDs in nursing home 70% females, Mean age (SD;</td>
<td>A room with one way mirror</td>
<td>Observational study (n = 20)</td>
<td>N/A</td>
<td>1. Observational data measuring PWD’s acceptance towards the robot</td>
<td>PWDs felt comfortable in interacting with robots in general. They touched the robot and liked the games. But PWDs showed difficulty in using the robot</td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>Author, year, Country</td>
<td>Participants</td>
<td>Setting</td>
<td>Study design (sample size)</td>
<td>Multi-functional robot</td>
<td>Control</td>
<td>Measures *</td>
<td>Finding</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------</td>
<td>---------</td>
<td>---------------------------</td>
<td>-----------------------</td>
<td>---------</td>
<td>------------</td>
<td>--------</td>
<td>------------------</td>
</tr>
<tr>
<td>(Nakamura et al., 2021) Japan</td>
<td>PWDs in nursing homes 78% females; Mean age (SD) (range): 89.33 (5.66) (81–98); MMSE range: 6–27/30</td>
<td>A community centre</td>
<td>Observational study (n = 15)</td>
<td>Two version of Pepper were used for 7 days: - Mobile Pepper that can be teleoperated to move and approach the participants -Stationary Pepper was placed in an easily accessible location</td>
<td>N/A</td>
<td>1. Observed interactions with the robots (frequency, behaviors)</td>
<td>-60% PWD (9/15) engaged verbally with the mobile robot, which was twice more than that of stationary robot -PWD with poorer cognition showed less verbal engagement with the moving robot, compared to those with better cognition.</td>
<td>Lower</td>
</tr>
<tr>
<td>(Schramek et al., 2021) Germany</td>
<td>People with dementia living at home 75% females Mean age (range): 71 (52–83); Majority with moderate dementia Caregivers Nil</td>
<td>People with dementia’s homes</td>
<td>Individual interviews (N = 4 families)</td>
<td>Puppet ELISA; Weekly visits to the PWD’s home; individual facilitated session; Led by a trainer</td>
<td>N/A</td>
<td>Individual interview before the intervention and after</td>
<td>1. Generally accepted by people with dementia and caregivers; Enjoyed the tailor-made approach of robots’ visits; Enjoyed the entertainment; social contact; reduced the burden of carers 2. Limitation: relied heavily on the skills of the facilitator, costly</td>
<td>Lower</td>
</tr>
</tbody>
</table>

AI: The apathy inventory; APADEM-NH: Apathy Scale for Institutionalized Patients with Dementia Nursing Home version; CSDD: Cornell Scale for Depression in Dementia; MMSE: Mini-mental status examination; S-MMSE: Severe- Mini-mental status examination; NPI: The neuropsychiatric Inventory; QUALID: Quality of Life Scale
CG: caregiver; LTC: long term care; PWD: Person with dementia; Silbot, NAO, MARIO, Betty, Matilda, Sophie & Jack, Papero, GUDIE were the name of the multifunctional robots; PWD: Person with dementia
A under the control column represents that the study either does not have a control group or control condition.
* Instruments which are under the section of ‘measures’ that had been validated in dementia are highlighted in purple. Other non-highlighted information is either non-validated instruments or trial data (e.g., dropout rate)
|| All studies assessed short-term effectiveness. **Significant results highlighted in bold**

As the study was conducted in two phases, two sets of expert samples were recruited. Sample size and demographic information shown on table was calculated based on the overall sample recruited from the two phases.
† Alert interaction rate is the % rate at which PWD made responses to the robot’s attention-seeking action
§ Information acquisition rate is the % rate at which the robot had successfully acquired the information received from the PWD
Summary Multifunctional robots appeared generally acceptable to people with dementia from three high quality qualitative study and some lower quality observational studies but were less acceptable in other lower quality studies. People with dementia found them difficult to use. The most valued functions were music and dancing.

3.6.2. Effectiveness

One RCT and two non-randomized interventional studies examined the effectiveness of multifunctional robots. (D’Onofrio et al., 2019; Valentí Soler et al., 2015) All were from Europe (Ireland, UK, and Spain) and of lower quality. They assessed the immediate effect on cognition, quality of life and neuropsychiatric symptoms including depression and apathy and included 159 people with dementia (mean age 80, 72% females).

The only RCT (Valentí Soler et al., 2015) was designed as a pilot study and compared NAO with PARO and usual care. NAO was found more effective than usual care, but not PARO, in reducing apathy. But there was no effect on quality of life and overall neuropsychiatric symptoms, and cognition deteriorated with NAO use. A pre/post-test study (Valentí Soler et al., 2015) found NAO has an immediate effect on improving overall neuropsychiatric symptoms, but not on cognition and apathy. A pre/post-test study of MARIO (D’Onofrio et al., 2019) found no improvement in quality of life and depression after MARIO use.

Summary Only lower quality studies examined effectiveness and evidence was inconsistent about multifunctional robots’ effect on apathy and neuropsychiatric symptoms. There was no evidence of benefits for cognition and quality of life. No studies examined long-term effects of multifunctional robots.

4. Discussion

This systematic review provides comprehensive evidence on the feasibility, acceptability, and effectiveness of robot use for care of people with dementia. We identified four categories of socially assistive robots which were companion robots (pet companion robots and humanoid companion robots), homecare assistive robots, telepresence communication robots, and multifunctional robots. Companion robots have been most frequently studied. This review showed that using socially assistive robots is usually feasible and acceptable in long term care settings. However, all types of robots had limitations affecting user experience, which could prevent their widespread adoption in dementia care.

We found that PARO, a socially assistive robotic seal is generally feasible and acceptable to people with dementia, family caregivers and staff in LTC but heavy, noisy and there were costs of training as well as costs of PARO itself. Based on the current data from meta-analysis of available studies about effectiveness, which were all moderate and lower quality, we found no clear evidence that PARO’s improves, either immediately after the intervention or in the long-term cognition, overall neuropsychiatric symptoms, including agitation, anxiety, depression, and apathy, or quality of life. Conclusion about the effectiveness of PARO is still limited because of paucity of higher quality studies in the field.

Unlike PARO, the feasibility and acceptability of companion robotic animals is mixed and unclear. There are several possible explanations for the differing results between PARO and robotic animals. Most participants in robotic animals’ studies had severe dementia, which is associated with poor responsiveness towards robots (Law et al., 2019) in addition, half of the quantitative studies on robotic animals used real animals as their control group, which may make the robotic animals seem less acceptable, compared to the usual care or plush toy controls typically used in the PARO studies. Finally, the novel appearance of a seal-like robot PARO, as compared to a more familiar dog or cat, may have elicited more interest. Interestingly, two recent studies with moderate quality reported that the use of augmented virtual reality function could further enhance the user experience of animal robots.

Among all the studies we reviewed, none had examined what can be added within a robotic session to further enhance the human-robot interaction, so this is a potentially important future research direction. There was no clear evidence about effectiveness of robotic animals as only one lower quality interventional study was available, showing effect on quality of life but not agitation, so more high quality studies are therefore recommended.

Four studies examined humanoid companion robots. People with dementia enjoy the interaction with these robots as, unlike animal robots, they are usually designed to speak but there are problems with speech recognition. Robot designers could improve conversational abilities of humanoid companion robots by improving speech recognition and dialogue system. Also, two out of four studies found participants preferred humanoid robots more closely resembling humans (human voices and movement, and smooth facial expression). Future research should further investigate whether the level of human resemblance could be important in acceptability of humanoid companion robot. Although the only RCT found no evidence of benefits of neuropsychiatric symptoms and quality of life, higher quality RCT are needed.

Multifunctional robots were examined in 15 studies and, although they were generally acceptable and feasible for people with dementia, they were difficult to use because of the interface and ergonomics. This could be because these robots were originally not designed specifically for people with dementia. Robot developers could improve the designs by considering the needs of people with dementia, such as having a large screen with adjustable tilting level, simpler instructions and speech, and a loud clear voice.

Telepresence communication robots and homecare assistive robots were each examined in only three studies. We found that telepresence communication robots were generally acceptable, but susceptible to internet connection problems, as they all required a consistent Wi-Fi signal. Our results aligned with a previous integrative review reporting acceptability and similar limitations (Moyle et al., 2017a). We also found limitations related to caregivers’ difficulties in operating the robots due to technical and time demands. Robot developers could improve interface designs to simplify using them, and future studies could consider examining time saved.

Evidence for homecare robots is limited as only two types of robots have been studied by four papers so far. There was mixed evidence on the feasibility and acceptability of homecare assistive robots, but in general people with dementia expressed more negative views than their family members. This was largely due to people with dementia believing they were sufficiently independent and not requiring the type of assistance offered by the robot. Refusal of daily care assistance is common among people with dementia (Volier et al., 2007) and being assisted by a ‘robotic machine’ may be more or may be less acceptable to some. Acceptability of homecare robots may improve if they are conceptualised as a benefit and a means for maintaining independence rather than as sign of a person’s deteriorating ability. Our review also showed that the non-humanoid feature, technical challenges, cost of homecare robots and their limited capabilities in complex homecare tasks were found to be barriers to their use. We were surprised to find that robot size, privacy or maintenance issues were not identified as barriers even though these issues have frequently been discussed in the literature on homecare assistive robots. (Robinson et al., 2014) This may reflect that all included studies of homecare robots were conducted in laboratory settings, so may not transfer to daily life for people with dementia living at home.

From our review, several ethical concerns were raised as limitations of robot use in dementia, including use of pet companion robots was infantilisation of people with dementia, intrusions on privacy from teleoperated communication robots and reduced human contacts by substituting from all types of robots. Cost of robots presented as a continual limitation throughout our findings, raising an ethical concern related to equality of access of robot use on people with different socioeconomic status. A recent study surveying ethical perceptions of robot
use on dementia identified similar concerns (Bradwell et al., 2020), although concluded that these ethical issues did not weigh strongly enough to act as barriers to successful real-world implementation. No ethical issues, except costs, were reported as a reason for not buying a robot. Future research should investigate what limitations of robots would influence real world implementation. Additionally, although some researchers have referred to the ‘deception’ of using a companion robot (Sharkey, 2014; Sparrow, 2002; Sparrow and Sparrow, 2006) to influence a person with dementia to feel socially connected, we did not find that deception was identified as a limitation by stakeholders. This may reflect that this may not concern them as long as interactions were beneficial to the people with dementia.

4.1. Strength and Limitations

This is the first systematic review which categorizes all types of socially assistive robots in dementia care, investigates the existing evidence for their feasibility and acceptability and considers both immediate and longer-term effectiveness. It is the most comprehensive review to date, as it included all study designs conducted from inception, had no language limits, and searched several subject databases. Strength of this review also includes having two independent raters for screening, quality appraisal and data extraction, prioritizing higher quality studies in synthesis and excluding findings from non-validated outcome measures.

We also used a single study quality assessment tool (MMAT) when assessing study quality. This allowed us to make comparison across studies with different study designs, making prioritization of studies with higher study quality possible. Our findings on effectiveness of PARO contrasts with a previous meta-analysis (Leng et al., 2019) which showed immediate effect on agitation and depression but unlike our study included results from unvalidated instruments. Like our study they did not find effects on cognition and quality of life. Our differing findings is due to our study only including studies which used validated outcome measures, thereby improving the strength of our conclusion.

The study however has limitations. First, there were only five moderate quality RCTs, and we could not answer many questions about effectiveness of different robots. Second, all studies were from high-income countries and thus there is no evidence about low- and middle-income countries. Third, we did not find any outcomes for caregivers when assessing effectiveness. Fourth, most participants were care homes residents, so our findings are less generalizable to community-dwelling population (Harrison et al., 2019). Fifth, in our meta-analyses, we were unable to perform cluster adjustment for the one cluster RCT (Joranson et al., 2015a, 2016c) due to data unavailability. This may increase the effect size, but, as we did not find any significant results on any outcomes in our meta-analyses, it did not affect our overall conclusion about PARO. Similarly, in a cross-over study (Moyle et al., 2013b) included in our meta-analysis, we used data for both study periods as the way the data presented did not allow us to perform analysis which includes the first period only. As the sample size of this paper is small and the result negative, we believe this did not affect our meta-analyses’ conclusion.

4.2. Implications

There is a body of opinion suggesting that robots are one of a range of practical solutions to address the increasing requirements and daily challenges of people with dementia and those who care for them as numbers increase. Further optimisation of these solutions will likely require investment in user experience and impact research and new developments in interactive robotic technology. Therefore, close collaborations between families, clinicians, social and behavioural scientists, and engineers should be encouraged at all stages of research and development to facilitate feasibility and acceptability of socially assistive robots.

There are several implications of our findings for future research. We need to become clearer about what constitutes meaningful outcomes of effectiveness in dementia care and apply this knowledge to our assessments of robot interventions. While companion robots did not show short- and long-term effects on cognition, neuropsychiatric symptoms, and quality of life, focusing only on them can miss other outcomes. To date, enjoyment and engagement while using socially assistive robots have been reported as an indicator of feasibility and acceptability of robotic devices but wellbeing (as opposed to health-related quality of life) has not been considered in full trials. There were no studies of homecare assistive robots at home and their usefulness for everyday tasks remains unknown. Future research should explore in greater detail what happens during interaction between the user, the robot and anyone else present. This might provide new insights into effectiveness and immediate benefits of robot use. Also many studies were predominately qualitative or small observational studies without control groups and used non-validated outcome instruments. Future studies should use high quality designs to establish further evidence with active control groups to find out if companion robots are as good as or superior to plush toys or real animals.

Our findings also suggest that adoption of robots, if found to be effective, may require changes in mindset and skills, with the associated additional demands on time and effort. For example, homecare assistive robots may need to be reframed to increase independence rather than undermine it. Caregivers and staff will need to allocate time to learn how to operate robots and support people with dementia in accessing them.

5. Conclusion

Overall, this review suggests that using robots in dementia care can be feasible and acceptable. People with dementia, family caregivers and staffs enjoyed the use of robots, but all types of robots have limitations which can potentially prevent widespread adoption. Robot designers and clinicians could overcome some of these limitations by improving the design of robots (to improve ease of use) and making sure the household are well-supported in case of any technical challenges. Based on the current data about effectiveness, there is no clear evidence that robots improved cognition, neuropsychiatric symptoms, or quality of life of people with dementia. But a lack of evidence is not evidence of lack of effectiveness, especially considering the dearth of higher quality papers in the field. We recommend that future research should use high quality designs with well-validated outcome measurements to establish evidence of effectiveness and also assess long-term effectiveness. We recommend the inclusion of measures of cost-effectiveness as cost is a major concern for stakeholders.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

GL conceptualised the study, CY performed the literature search, CY and LS made decisions about whether articles fulfilled the inclusion criteria and rated the quality of articles with input from AS and GL as needed. CY and AS performed meta-analyses. GL and CY wrote the first draft. All authors read the draft and critically revised it for important content and approved the decision to publish. GL is the guarantor and accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish.

Acknowledgements

GL and AS are supported by UCLH National Institute for Health Research BRC. GL is supported by the National Institute for Health Research.
Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ajrr.2022.101633.

References


