BMJ Open

A Systematic Review on the Use of Socially Assistive Robot Technology in Elderly Care

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-018815
Article Type:	Research
Date Submitted by the Author:	26-Jul-2017
Complete List of Authors:	Abdi, Jordan; Imperial College London, Al-Hindawi, Ahmed; Chelsea and Westminster Hospital, Magill Department of Anaesthesia Ng, Tiffany; Chelsea and Westminster Hospital, Magill Department of Anaesthesia Vizcaychipi, M; Imperial College London, Magill Deaprtment Anaesthesia
Primary Subject Heading:	Geriatric medicine
Secondary Subject Heading:	Sociology, Public health, Mental health
Keywords:	Socially Assistive Robots, Elderly, Robots, GERIATRIC MEDICINE, SOCIAL MEDICINE

A Systematic Review on the Use of Socially Assistive Robot Technology in Elderly Care

Jordan Abdi[1], Ahmed Al-Hindawi[2], Tiffany Ng[2], Marcela P Vizcaychipi[2]

[1]Department of Medicine, Imperial College London, Kensington, London SW7 2AZ

[2]Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH

Author note:

Jordan Abdi (Abdi J), Undergraduate student, Department of Medicine, Imperial College London

Ahmed Al-Hindawi (Al-Hindawi A), Intensive Care Registrar, Magill Research Fellow, Magill Department of Anaesthesia, Chelsea and Westminster Hospital

Tiffany Ng (Ng T), Doctor, Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH.

Marcela Vizcaychipi (Vizcaychipi MP), Consultant in Anaesthesia and Intensive Care in the Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH

Address Correspondence: Jordan Abdi, Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH.

Electronic address: ja1612@ic.ac.uk

Use of Assistive Robot Technology in Elderly Care

STRUCTURED ABSTRACT

OBJECTIVE: The demand for geriatric care is expected to increase with an elderly population that is set to more than double by 2050 worldwide. This trend is paired with a fall in the number of younger people able to support the senior members of society, posing several problems to how we continue delivering high quality care. Socially Assistive Robots (SAR) promise to reform elderly care, and many have imagined potential roles they could serve. This review qualitatively examines the literature on the use of SAR in elderly care and aims to establish the roles SAR may play in the future.

DESIGN: Systematic Review

DATA SOURCES: Systematic search of CINAHL, Cochrane Library, EMBASE, MEDLINE, PsychINFO and SCOPUS databases was conducted, complemented with a free search using Google Scholar and reference harvesting. All publications went through a selection process, based on pre-defined selection criteria, which involved sequentially reviewing the title, abstract and full text of the publication. No limitations regarding date of publication were imposed and only English publications were taken into account.

ELIGIBILITY CRITERIA: The inclusion criteria consist of elderly participants, any elderly healthcare facility, humanoid and pet robots, and all social interaction types with the robot. Exclusions were acceptability studies, technical reports of robots and publications surrounding physically or surgically assistive robots.

RESULTS: In total, 51 final publications were included in the review, describing 27 studies and including 908 participants and 9 robots. Five roles of SAR were identified: Affective Therapy, Cognitive Training, Social Facilitator, Companionship and Physiological Therapy.

CONCLUSIONS: Although many positive outcomes were reported, a large proportion of the studies have methodological issues which limit the utility of the results. Nonetheless the reported value of SAR in elderly care does warrant further investigation, and future studies should endeavour to validate the roles demonstrated in this review.

SYSTEMATIC REVIEW REGISTRATION - NIHR 58672

Abstract Word Count: 298 Words

Keywords: Socially assistive robots, elderly, robots, geriatric care, social care

Strengths and Limitations

- This review had a very clear and focused question, with two main search search concepts: Socially Assistive Robots and Geriatric Care. This allowed the review to encompass all aspects of geriatric care, whilst maintaining defined parameters.
- This review's methodology used clear and comprehensive inclusion and exclusion criteria, minimalising ambiguity in the screening phases.
- The use of SAR in elderly care is still a small field and despite using wide search criteria, only a handful of relevant studies were found. This dramatically restricts the quantity and quality of information available to validate or dismiss the utility of SAR.
- The roles were created retrospectively by this review, as part of a discovery process on extracting data from the final set of studies. Whilst they have utility in evaluating the state of the field and providing defined expectations for the technology, they have generalised sets of studies that are very different in quality, design and sometimes outcome. As a consequence, this may mislead the actual weight of data in the respective roles.
- This review ran the inadvertent risk of excluding relevant papers in the screening phase. Although high concordance between the reviewers was reported, the large volume of studies that had to be reviewed invites the possibility that relevant publications were excluded. The reason for the high exclusion rate was the broad search criteria identified irrelevant robot interventions, such as surgical robots or telecommunication devices. It is unlikely, however, that an additional study would have changed the conclusions the review came to.

Introduction

The global population is undergoing a demographic shift as life expectancy grows and the post-war baby boom generation enters retirement. The implications on resource allocation will impact the delivery of elderly care. As of 2015(1), 21% of Western Europe's population were over the age of 60, and this is expected to rise to 33% by 2030. By 2050, there are expected to be more people over the age of 60 globally than under 15, reaching a total population of 2.1 billion compared to 901 million in 2015. This is compounded by a proportional decrease in the number of social and health care providers shouldering this increased burden. In 2015, 7 workers were allocated for every elderly person globally, but this is projected to fall to 4.9 in 15-years(1). Moreover, the situation is magnified in Europe by an accelerated ageing population with 3.5 workers for every elderly person and set to fall further to 2.4 by 2030. This shift in societal proportions will place new pressures on elderly care.

Loneliness is a consequence of social, psychological and personal factors. Over half of those over the age of 75 live alone(2) and 17% of older people see family, friends or neighbours less than once a week(3). A recent meta-analysis(4) showed that the impact of loneliness and isolation carries the same mortality risk as smoking 15 cigarettes a day. This is compounded by the fact that social care is labour intensive industry in a world with a proportionally shrinking workforce.

Throughout many industries the 'robot revolution' promises to solve this growing personnel shortage. At present, physically or surgically assistive robots dominate the healthcare sector's robot usage, including: (i) increasingly sophisticated wheelchairs transforming the limitations imposed on paraplegics; (ii) robotic limbs redefining amputee capabilities; and (iii) robotic surgeons revolutionising how and where surgery can be

performed. Nonetheless, physically assistive robots do not combat the increasing mental health burden recognised in the elderly population. The list of predicaments facing an elderly person extends beyond the physical. In response to this challenge, the concept of socially assistive robots is gaining headway. These are robots adept at completing a complex series of physical tasks with the addition of a social interface capable of convincing a user that the robot is a social interaction partner(5).

Socially assistive robots have been categorised into 2 operational groups: 1) service robots and 2) companion robots. Service robots are tasked with aiding activities of daily living(6). Companion robots, by contrast, are more generally associated with improving the psychological status and overall well-being of its users; examples include Sony's AIBO(7) and Paro(8). Despite much of the hype, the utilisation of this technology in elderly care is not completely ascertained.

The aim of this systematic review is to establish the roles of Socially Assistive Robots (SAR) in elderly care. This novel perspective on the burgeoning technological field addresses what expectations are reasonable in the coming decades and sketch a future direction for advances in this field. Through examining the existing literature and qualitatively analysing them, this review identifies studies that have demonstrated utility of SAR. Our findings synthesize a holistic picture encompassing aspects of social and physiological wellbeing that can be offered to the elderly population.

Methodology

In March 2016, a systematic review was conducted in accordance with the principles of the Cochrane Handbook for Systematic Reviews of Interventions (9). This systematic review's registration number is NIHR 58672. The CINAHL, Cochrane Library, EMBASE, MEDLINE, PsychINFO and SCOPUS databases were searched for publications regarding SAR in elderly care. No limitations regarding date of publication were imposed and only English publications were taken into account.

To ensure that all relevant publications were included, and to prevent subjective biases from omitting relevant material, 2 reviewers independently screened the publications. The screening process involved 3 assessment stages, whereby the title, abstract and full text were sequentially reviewed. To ensure no publications were missed, additional publications were selected through a free search (Google Scholar) and from reference lists of selected publications and relevant reviews.

The database search query was composed of 2 search concepts. The first was the intervention (SAR) and the second was the context (elderly care). The aforementioned databases were searched using these concepts and incorporated both free terms and Medical Subject Headings (MeSH) terms, however where appropriate, the database specific thesaurus equivalent was used. The free words used for the intervention were "service robot*", "therapeutic robot*" and "socially assistive robot*", and their associated MeSH terms were "Robotics" and "Artificial Intelligence", including their subheadings. To avoid excluding suitable publications, the names of specific robot systems were also searched for, namely "AIBO", "Care-o-bot", "CERO", "Feelix", "Huggable", "iCat", "Ifbot", "Matilda", "NAO", "NeCoRo", "PaPeRo", "Paro", "Pepper", "Robocare" and "Sparky". The free words used for the context were "elder*", "senior*", "older person*", "old people" and

"dementia", and their associated MeSH term was "Aged, 80 and over". The use of the asterisk (*) enables the word to be treated as a prefix. For example, "elder*" will represent "elderly" and "eldercare" amongst others.

In addition to this systematic search, two further means of acquiring relevant publications were used. Existing publications and reviews of the field of SAR may contain relevant publications not identified in the database search and were screened and added accordingly to the pool of potentially eligible publications. Additionally, a free search (Google Scholar), was performed to identify any remaining publications.

Technical reports of robots and publications surrounding physically or surgically assistive robots were excluded. Feasibility or acceptability studies were excluded, unless the study also assessed a potential utility of SAR. No publication was excluded on the grounds of publication date.

The central criterion by which publications were selected was whether the publication had assessed the utility of SAR in the elderly population, above and beyond simply assessing its acceptability. All publications collected during the database search, free search and reference harvesting were judged for relevance by the 2 reviewers in a series of assessment rounds independently. The first was based on title, then abstract and finally full text. A 3-point scale (0 = Not relevant, 1 = Possibly relevant, 2 = Very relevant) was used by the reviewers to score the publications, and those with a combined score of 2 would make it through to the next round of scoring. All publications with a total score of 0 were excluded and would not proceed to the following round. A publication with a combined score of 1 indicated a disagreement between the reviewers, and would be resolved through discussion. At the end of the full text screening round, a final set of publications to be included into the review was acquired. Cohen's kappa coefficient was calculated to

ascertain the agreement between the reviewers in the title, abstract and full text screening phases.

Since the field of Socially Assistive Robotics is very much in its infancy, the nature of many of the studies can be expected to be small and exploratory. However, they do offer novel insight into potential applications of SAR in a variety of elderly care institutions and for this reason, no publication was excluded on the grounds of methodological quality.

The data extraction form for this review was designed in line with the PICO approach (Participants, Intervention, Comparator and Outcomes). This process was conducted by one reviewer to ensure consistent extraction of all studies.

During the data extraction phase, studies were categorised by the role of the robot in the study. The categories were generated by the retrospectively and were not pre-defined or directly referenced in the original studies themselves.

Duplicate reports of the same study can overstate the significance of certain interventions and can draw greater than warranted attention to its findings. Such duplications may present in different journals, papers or conference proceedings, and may each focus on different outcome measures or include a follow up data point. Identifying duplicates is an exercise of cross-assessing multiple reports, and ensures that the principles and findings of each study get fair representation. To achieve this, the final set of reports were collated into “study groups”. The data extraction process was conducted on the most comprehensive report of a given study.

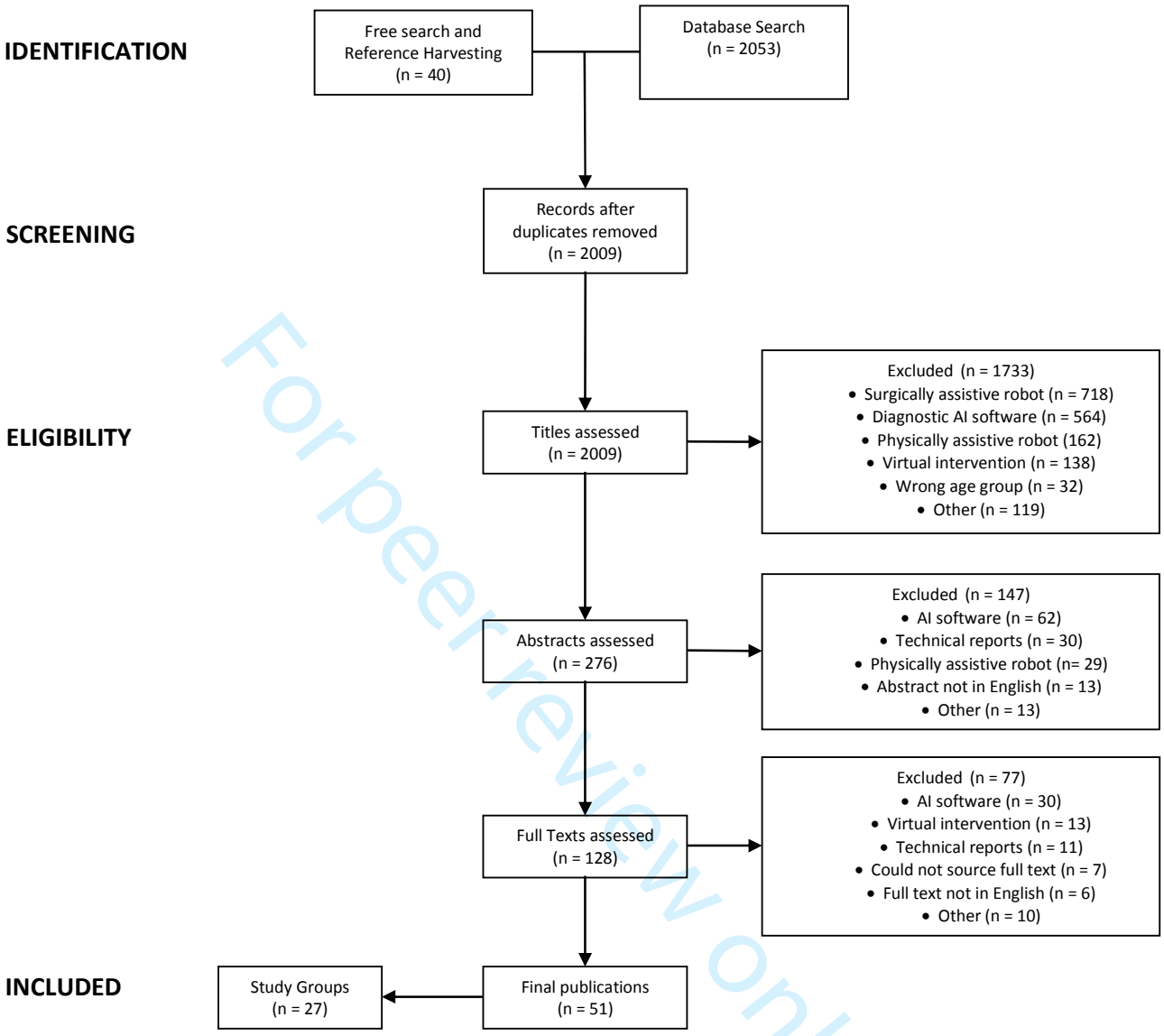

Figure 1: Schematic flow diagram of the review process

RESULTS

Search Results

The database search yielded 2053 publications and a further 40 were included from reference harvesting and the free search. Duplicate publications were removed (n=84) and following 3 screening phases, 51 publications were eligible and included in the review. Once duplicate reports were collated, a total of 27 original studies were identified and subject to detailed review.

The inter-rater agreement between the reviewers were calculated to be 0.91 for the title screen, 0.64 for the abstract screen and 0.89 the final report; demonstrating very good, good and very good correlation between the reviewers respectively according to Cohen's Kappa coeffecient (10).

Figure 1 outlines a PRISMA schematic flow diagram of the review process and reasons for exclusion (11).

Participants and Settings

Across the studies, 908 participants were included, however, due to inconsistent reporting, overall age and gender information are not available. All participants were considered elderly (aged > 60 years) with a single exception. The number of participants included in any given study varied from 3 to 110 subjects. In the 18 studies that reported gender information (comprising 594 participants), 78% of the participants were women. A large proportion of studies were conducted in Japan (n = 10; 178 participants) and the US (n = 6; 98 participants), and most of the studies were conducted in a nursing home (n = 15; 593 participants). Details of robot systems reviewed can be found in Table 2.

Identified Roles of Socially Assistive Robots

Eligible studies were organised into sets by the role assumed by SAR. Five roles were identified: *Affective Therapy*, *Cognitive Training*, *Social Facilitator*, *Companionship* and *Physiological Therapy*. Specific details of the studies below, such as assessment tools or subject demography, are described in Table 1.

Affective therapy

Eleven studies evaluated the effect SAR can have in improving the general mood and well-being of elderly participants, or its ability to overcome episodes of mood disturbance, a role this review has collectively termed *Affective Therapy*.

In 2 studies, interactions with SAR improved subject mood-scale scores. A randomised controlled crossover trial(12) of 18 subjects with dementia in Australia assessed the emotional effect of group activity with Paro. Subjects were randomised into the Paro intervention group or an interactive reading group. Quality of life and mood were assessed at baseline, 5-weeks and 10-weeks. The study reported that the Paro intervention improved quality of life and pleasure scores, whilst reducing anxiety and sadness scores.

A larger study (13) of 53 subjects with dementia across 10 nursing homes in Norway assessed the effects of Paro group activity on symptoms of agitation and depression. In this cluster-randomised controlled trial, subjects were randomised into group activity with Paro or receive treatment as usual (control). Cognition was assessed at baseline, 12-weeks and at 3-months post-intervention. The intervention group's mean scores decreased significantly for agitation ($p < 0.05$) and depression ($p < 0.05$) across all time points compared to controls.

In a more complex 4-month ABAB crossover study(14) with 71 elderly subjects in the Netherlands, the outcomes of Paro interventions were evaluated. The two active phases, interspersed with control phases, compared two types of Paro interventions: therapeutic intervention (Paro introduced at times when subject was distressed) and care support intervention (Paro introduced to facilitate activities of daily living). Only the therapeutic intervention showed a significant improvement in the mood score ($p < 0.01$).

In one Japanese pilot study(15), mood-scale scores were improved but this was not sustained at follow up. The group interactions of 26 subjects with Paro found significant improvements in mood scores for weeks 2-5 ($p < 0.05$). Although this was a temporary effect, nurses did comment on improved sociability between subjects and an improvement in the subjects' overall enjoyment of the service centre.

The interest level in Paro was studied in a pilot study(16) of 14 elderly subjects in Japan. The psychological effect of group activity with Paro over 5-years was assessed and interest in Paro over the long-term was maintained. Accordingly, it may be possible that longer-term interactions with Paro could mitigate the temporary effect seen in the previous study.

In 2 studies, no improvements in quality of life or mood scores were found. The pilot study(17) of 4 elderly subjects with dementia in Sweden assessed how weekly interactions with JustoCat over a 7-week period would affect self-reported quality of life and agitation. Contrary to the other studies, no changes in baseline compared to post-intervention scores were reported in any of the participants. Similarly, no changes in any general mood or depression scores in a randomised controlled trial(8) of 23 subjects in Japan were found. In this trial, the effect of group interactions with Paro and placebo Paro (robot switched off) on were unremarkable.

Apart from mood, a pilot study(18) in Japan found positive behavioural expressions of 30 subjects with dementia when evaluating the subjects' interactions with Paro compared to a stuffed lion. Over a 3-6-month period with 2 separate 15-minute interactions, subjects talked more frequently to Paro ($p < 0.05$) and showed more positive emotional expressions with Paro ($p < 0.01$) than with the stuffed lion.

Four studies produced mixed results. A randomised controlled trial(19) of 100 subjects in nursing homes across Denmark evaluated the effect of Paro sessions on sleeping patterns and psychiatric well-being of elderly people. Subjects were randomised into visit groups; interacting with Paro, a living dog or soft toy cat. Sleep data was collected and scales for cognitive state, independence and depression were assessed. The study found that visit type did not affect cognitive state, independence or depression scores and did not affect sleep quality. Over the course of the 6-week study, cognitive state and independence scores worsened across all the groups ($p < 0.05$) however depressive scores improved compared to baseline scores ($p < 0.05$).

A pilot crossover study(20) of 9 subjects with dementia in the US evaluated the psychological benefits of interactions with NeCoRo, compared to a toy cat. Subjects were randomised into each group, and switched onto the other intervention the next day. Agitation, mood and engagement was assessed at baseline and during each session. Compared to baseline scores, agitation scores were only significantly decreased in the toy cat ($p < 0.05$), whereas NeCoRo yielded significantly improved scores of pleasure and interest ($p < 0.01$ and $p < 0.05$, respectively).

A 2-phase crossover trial(21) of 20 and 17 subjects with dementia, in the respective phases in Spain evaluated whether SAR could improve behaviour, mood or quality of life in subjects. Subjects were assessed for cognitive deterioration, cognitive state, apathy and

quality of life at baseline and at a follow-up assessment after the end of each phase. In phase 1, deterioration scores increased, however scores for irritability decreased. In phase 2, deterioration scores increased.

Cognitive Training

Six studies assessed whether SAR can improve aspects of cognition, such as working memory or executive function, and as such this review has termed this set *Cognitive Training*.

A randomised controlled trial(22) of 34 healthy female subjects in Japan assessed the effects of living with Nodding Kabochan on the subjects' cognitive function. Only subjects receiving the Nodding Kabochan demonstrated improved cognitive function score ($p < 0.01$), higher judgement ($p < 0.05$) and verbal memory scores ($p < 0.05$), lower fatigue scores ($p < 0.05$), higher scores of motivation ($p < 0.01$) and decreased saliva cortisol levels ($p < 0.05$) at 8 weeks compared to control (non-functional Nodding Kabochan).

A pilot study(23) of 3 female subjects with dementia in the US assessed whether Bandit could improve performance in a musical game and in overall cognitive function. The study concluded that robot encouragement improves quiz performance.

Brain imaging was utilised in a randomised controlled trial(24) of 71 healthy subjects in South Korea. The study investigated if robot-assisted cognitive training altered cortical thickness in brains of elderly participants. Subjects were randomised into 3 arms: (1) robot-assisted group training using Silbot and Mero; (2) traditional intervention training, using computer software; or (3) non-intervention arm. The study showed attenuation of cortical thinning on MRI in both intervention groups ($p < 0.05$), and estimated it would take 15.3-months for intervention groups to reach the same level of cortical thinning as controls. The

intervention groups showed greater improvement in the executive function scores than control group ($p < 0.001$). In the general cognitive and visual memory tasks, the traditional intervention group had greater improvement than in the robot group. The robot group did not outperform the traditional group on any neuropsychological tests.

Similarly a pilot study(25) of 14 subjects with dementia in Japan used electrophysiological analysis in its assessment of SAR. This study investigated the neuropsychological influence of Paro within an interactive group setting. The authors analysed electroencephalogram recordings and subjective questionnaires and found an increase in cortical neuronal activity in 7 participants, particularly in participants who liked Paro.

Cognitive training in the form of a game took place in a pilot study(26) of 11 subjects with dementia in Japan that assessed if group activity with AIBO affected the behaviour of subjects. Subjects were assessed on their behaviour towards the robot during card-memory or ball games across 5-days. The authors concluded that the frequency of talking and physical gestures towards the robot increased over the study period.

An improvement in cognitive function is not seen in a 2-phase block randomised controlled trial(21). This Spanish study involved 101 and 110 subjects with dementia, in the respective phases, and assessed the cognitive effects of group interactions with SAR. Compared to control group at follow up, Phase 1 showed a decrease in cognitive function scores in the NAO group ($p <0.05$). Interestingly, Phase 1 also evaluated other psychiatric parameters compared to the control. The results showed a decrease in apathy scores in NAO and Paro groups ($p < 0.05$ respectively), increased delusions in the NAO group ($p < 0.05$), increased irritability scores in Paro group ($p < 0.05$). Notably, there were no significant differences between NAO and Paro groups at follow-up. Phase 2 compared to control group

at follow-up, showed an increase in quality of life scores in the Paro group ($p < 0.05$), and increased hallucinations and irritability in both the Paro and dog groups ($p < 0.05$ respectively). The Paro group demonstrated increased disinhibition ($p < 0.05$) and decreased night-time behaviour disturbances ($p < 0.05$) when compared to the dog group at follow-up.

Social Facilitator

Five studies assessed the utility of SAR as facilitators for improved sociability between subjects or between subjects and other people. As such this review has titled this role *Social Facilitator*. All of these studies concluded that the respective SAR intervention improved sociability of participants.

A crossover study(27) of 23 subjects in the US assessed whether Paro could stimulate sociable behaviour in subjects. Subjects were grouped into sessions with Paro and its normal functionality, Paro switched off, or no object. The authors report that the group with switched-on Paro engaged in more social interactions than the group with Paro switched off; no data was provided.

The role of communication skills was investigated in a pilot study(28) of 12 subjects in a residential care facility in Taiwan which assessed the effect of Paro on group activity participation and communication between subjects. The study found, after the 4-week programme, a significant improvement in communication and interaction skills exhibited by subjects ($p < 0.05$) and an increase in activity participation ($p < 0.05$).

Similarly a pilot study(29) of 8 subjects with dementia in Japan evaluated whether SAR improved communication between subjects. This study assessed the effect of individual interactions with AIBO on the cognitive state, dementia score and behaviour exhibited by subjects. The results showed that there was an increase in communication with care staff

during and after the interaction, however, no other changes were observed in any of the other assessments of cognitive state, dementia, or behaviour.

A crossover study(30) of 18 female subjects with dementia in the US assessed the impact of AIBO on the sociability of the subject. The study concluded that although all visit types, with AIBO, a dog, or no object, stimulated social interaction by the subject, there were no significant differences in frequency of social behaviours exhibited by subjects between the visit types.

A similar set of results were found in a US pilot study(31) of 7 subjects with dementia, which assessed whether group activity with Paro would elicit social behaviour from subjects towards others. Subjects within a group were divided into primary users, the individual engaging with Paro at any one time, or non-primary users, everyone else in the group. The study concluded that there were no significant differences between behaviours at baseline compared to those noted at the final session. This is despite the increased social interaction found over the 7-week period between primary and non-primary users towards each other and towards staff.

Companionship

Three studies assessed the utility of SAR in overcoming the feeling of loneliness and social isolation in the elderly. These studies are collected into a set this review has titled the *Companionship* role. All 3 of the studies examining SAR in this role showed significant reductions in loneliness scores.

A randomised controlled trial(32) of 38 subjects in the US compared the ability of a dog or AIBO to treat loneliness. Subjects were randomised to have weekly sessions with a dog, AIBO, or no object (control). Subjects in the dog or AIBO group were significantly less lonely

than those in the control group at week 7 ($p < 0.05$ respectively) and both interventions scored highly in attachment compared to the control group. No significant differences were found between the dog and AIBO groups in the assessment of loneliness, or attachment.

A randomised controlled trial(33) of 34 subjects in New Zealand investigated the psychosocial effects of Paro activity on loneliness. Subjects were randomised into a Paro group or a control group that attended standard activities. Subjects in the Paro group had a significantly greater decrease in loneliness score at follow-up than the control group ($p < 0.05$).

A pilot study(7) of 11 subjects in Japan evaluated the effectiveness of AIBO in improving the quality of life of elderly participants. Between the first and last AIBO session, "emotional words", "amount of speech" and "satisfaction" scores significantly increased ($p < 0.05$). Mean loneliness scores after the session were significantly lower than those before the session ($p < 0.05$).

Physiological Therapy

Two studies investigated the effects of SAR on physiological markers, and as such this review titles this set *Physiological Therapy*.

A pilot study(34) of 21 subjects in New Zealand investigated the effect of interacting with Paro on blood pressure and heart rate. Subjects had a single 10-minute session with Paro where they were free to interact with the robot. Blood pressure and heart rate was recorded before (T1), immediately after (T2) and 5-minutes after (T3) the 10-minute interaction. Overall, no significant changes in blood pressure or heart rate were demonstrated, however the study decided to exclude 4 residents who did not interact or touch the robot. Subsequently, significant decreases in systolic blood pressure ($p < 0.05$)

from T1 to T2 were shown, and such decreases were sustained at T3 measurement. Similarly, significant decreases in diastolic blood pressure ($p < 0.05$) from T1 to T2 were shown, however this decrease was not sustained at T3. Between T1 and T3 heart rate significantly decreased ($p < 0.05$).

In another pilot study(35) of 12 subjects in Japan, the physiological and social effects of interacting with Paro were investigated. Compared to baseline readings, a significant increase in the ratio of urinary 17-ketosteroid:17-hydroxycorticosteroid ($p < 0.01$), by week 4 of Paro being introduced, was found. The authors suggest this confers an improved physiological reaction to stress. A confounder noted was an increase in social interactions with other residents ($p < 0.05$) by week 4, compared to baseline.

Quantitative Comparison

Several studies reported comparative quantitative data, by using the same or similar assessment scales to others within their role category. The data from these studies have been reproduced from the studies and are compiled in Tables 3-5. Five comparable studies were identified in the Affective Therapy, each using a mood scale to assess either anxiety or depression. Of these, 3 showed significant improvements in the mood scores either in the robot intervention group or in the follow-up score, depending on study design. Four comparable studies were identified in the Cognitive Training set of studies, each using a validated scale to assess cognitive function. Of these, 3 showed significant improvements in the cognitive scores. Depending on the study design, this was evident in either the robot intervention group or in the follow-up score. Finally, three studies with comparable data were identified in the Companionship set of studies, each of which used validated loneliness scales to assess the subjective sense of loneliness. Two of these studies showed significant

improvements in the loneliness scores in the robot intervention group or in the follow-up score. No comparative data was identified in the Social Facilitator or Physiological Therapy groups.

Use of Assistive Robot Technology in Elderly Care

Table 1: Characteristics of Selected Studies

Role	Ref.	Participants	Setting	Intervention	Duration	Measures	Outcome
Affective Therapy	(17)	4 subjects (2 men) aged 82-90, with dementia	Dementia care home, Sweden	Supervised one-on-one interaction with JustoCat. Pilot study.	1 session (Unknown time length)/ week for 7 weeks.	QUALID, CMAI and interview	1. No significant changes observed in scales
	(18)	30 subjects (19 with mild/ moderate dementia + 11 with severe dementia), mean age 84.9 years (mild/ moderate), 87.5 (severe)	Nursing care facility, resident's room, Japan	Supervised one-on-one interaction with Paro and Stuffed Lion. Pilot study.	1 session (~15 mins) for each intervention per subject, separated by 3-6 months	Observed behaviour seen in video-recording	*In both groups:* 1. Subjects talked more frequently to PARO 2. Showed more positive emotional expressions with PARO *In Mild/ moderate group only:* 1. Showed more negative emotional expressions with Lion 2. Frequencies of touching and stroking and frequencies of talking to staff member were higher with Lion *In Severe group only:* 1. Showed neutral expression more frequently with Lion
	(14)	71 subjects (14 men) with dementia in 2 groups: therapeutic intervention and care support intervention	Psychogeriatric care institutions, Netherlands	Supervised one-on-one interaction with Paro or no intervention. Paro either served as a therapeutic or a care support tool in two separate phases of the study. Crossover study.	5 sessions (~15 mins)/ month for 2 months; each month of therapy was interspersed with a control month. In the therapeutic arm only, additional sessions were given when patient was in distress.	IPPA and Coop/ Wonca after each interaction	1. Therapeutic-related interventions show an increase of IPPA scores by 2 points 2. Care support intervention showed no effect
	(13)	53 subjects (20 men) aged 62-95, with a cognitive impairment (MMSE < 25) or diagnosed dementia	Nursing Home, separate room, Norway	Supervised group interaction with Paro or TAU. Randomised controlled trial.	2 sessions (~30 mins)/ week for 12 weeks	Cognitive status, medication, BARS and Norwegian version of CSDD. Assessed before (T0), after (T1) and at 3-month follow-up (T2)	1. Reduction in agitation in Paro vs TAU from T0 – T2 2. Reduction in depression in Paro vs TAU from T0 – T2
	(12)	18 subjects, aged >65, with dementia	Nursing Home, Australia	Supervised group interaction with Paro or reading group. Randomised controlled trial.	3 sessions (~45 mins)/ week for 5 weeks	Modified QoLAD, RAID, AES, GDS, Revised Algase Wandering Scale–Nursing Home version and OERS	1. The Paro group had higher QOL-AD and OERS-Pleasure scores following the intervention 2. The Paro group had reduced OERS-Anxiety and OERS-Sadness scores following intervention
	(16), (36), (37)(38), (39), (40), (41), (42)	14 subjects (all female) aged 77-	Health Service	Free group interaction with Paro. Pilot study.	2 sessions (1 hour)/ week for 1 year (and a 5-year	Face scale, GDS and Nursing comments	1. A tendency to improve depression after 8 weeks 2. Improvement in mood

Use of Assistive Robot Technology in Elderly Care

		98, 1 subject without dementia	Facility, Japan		follow-up)		3. Patients did not lose their interest in the long-term
	(19)	100 subjects with a mean age of 85.5	Nursing Homes, Denmark	Supervised one-on-one interaction with Paro, Dog or Toy cat. Randomised controlled trial.	Two sessions (10 mins)/ week for 6 weeks	MMSE, GBS, GDS, CAM, Sleep data and BMI	1. Greater interaction with Paro and Dog compared to toy 2. MMSE and GBS scores worsened over study period in all groups 3. GDS scores improved over study in all groups
	(20)	9 subjects (all female) aged 83-98, with dementia	Nursing Home, US	Supervised one-on-one interaction with NeCoRo and Toy cat. Crossover study.	1 session (10 mins) for each intervention	ABMI, LMBS and observations	1. Both cats maintained participant's interest 2. Significant pleasure and interest increase whilst playing with NeCoRo
	(15), (43), (44)	26 subjects (all female) aged 73-93, some subjects had dementia	Day service centre, Japan	Free group interaction with Paro. Pilot study.	3 sessions (~45 mins)/ week for 5 weeks	Summarised POMS, Burnout scale for nursing staff, nursing staff comments	1. Significant improvement in POMS scores 2. Positive social and psychological effects
	(8), (45), (46), (47), (48), (49)	23 subjects (6 men) mean age 85	Health Service Facility, Japan	Free group interaction with Paro or placebo Paro. Randomised controlled trial.	4 sessions (1 hour)/ week for 4 weeks	POMS, Face scale, Urinary tests and Nursing comments	1. Improvement in moods and reduction in depression and dejection levels in both groups 2. Urinary results suggest Paro interaction reduces stress
	(21)	Phase 1: 20 subjects (10 men), mean age 77.9 Phase 2: 17 subjects (8 men), mean age 79 All subjects were diagnosed with dementia	Day Care Centre, Spain	Phase 1: Supervised group therapy (Cognitive and Physical) with Nao. Phase 2: Supervised group therapy (Cognitive and Physical) with Paro. Crossover study.	2 sessions (30-40mins)/ week for 3 months	GLDS, sMMSE, MMSE, NPI and AI	*Phase 1:* 1. Increase in GDS scores 2. Significant decrease was seen in NPI-irritability/lability scores and total NPI scores *Phase 2:* 1. Increase in GDS scores
Cognitive Training	(22)	34 subjects (all female), aged > 65 years, living alone	Participant's home, Japan	Living with Nodding Kabochan or control robot (same design as Nodding Kabochan, but cannot talk or nod). Randomised controlled trial.	8 weeks	Questionnaires, BMI, Cognitive tests, APG, and blood and saliva samples	1. Cognitive scores (MMSE + components of Cognistat) were improved in Nodding Kabochan group 2. Saliva cortisol level was decreased in Nodding Kabochan group 3. Higher reports of loss of fatigue, enhancement of motivation and healing in Nodding Kabochan group

Use of Assistive Robot Technology in Elderly Care

(21)	Phase 1: 101 subjects (13 men), mean age 84.7 Phase 2: 110 subjects (11 men), mean age 84.7 All subjects were diagnosed with dementia	Nursing Home, Spain	Phase 1: Supervised group therapy (Cognitive, Musical and Physical) with Paro or Nao or TAU. Randomised controlled trial. Phase 2: Supervised group therapy (Cognitive, Musical and physical) with Paro or Dog or TAU. Randomised controlled trial.	2 sessions (30-40mins)/ week for 3 months	GLDS, sMMSE, MMSE, NPI, APADEM-NH and the QUALID	*Phase 1:* 1. Decreased apathy in NAO and Paro groups 2. Increased delusions in the NAO group 3. Increased irritability in both robot groups 4. Decrease in scores on the MMSE, but not the sMMSE, in the NAO group 5. There were no significant differences between NAO and Paro groups *Phase 2:* 1. Increase QUALID scores in the Paro group compared to the TAU group 2. Increased hallucinations and irritability in both the Paro and Dog groups compared to the TAU group 3. Increased disinhibition in Paro group compared to Dog group 4. Decreased night-time behaviour disturbances in the Paro group compared to Dog group 1
(24)	71 healthy subjects, aged >60, based in community	Assessment centre, South Korea	Supervised group interaction with either Silbot and Mero robots (robot cognitive training) or on-screen quiz (traditional cognitive training) or received no cognitive training (control). Randomised controlled trial.	5 sessions (90 mins)/ week for 12 weeks	MRI, Neuropsychometric tests and Alzheimer's Disease Assessment Scale	1. An attenuation of cortical thinning in both intervention groups 2. Robot therapy showed significantly reduced cortical thinning in the right and left anterior cingulate cortices and small areas of right inferior temporal cortex compared to traditional intervention 3. Global topological organization of white matter corticocortical networks was decreased in the control group and the rate of decrease was significantly less in both the intervention groups 4. Robot therapy had greater nodal strength in the left rectus gyrus 5. The intervention groups showed greater improvement in the executive function 6. In the general cognitive and visual memory tasks, the traditional intervention group had greater improvement than in the robot group 7. The robot group did not outperform the traditional group on any neuropsychological test
(23), (50), (51), (52)	3 subjects (all female) aged >70 with dementia (some reports say 4 subjects, with 1 male)	Care facility, US	Individual interaction (musical, cognitive game) with Bandit (compared to an on-screen simulation of Bandit in some reports). Pilot study.	One session (20 mins)/ week for 12 months	sMMSE, Response time, Correctness evaluation and Questionnaire	1. Robot encouragement improved response time

Use of Assistive Robot Technology in Elderly Care

	(26)	11 subjects with dementia	Nursing home, Japan	Interaction with AIBO, either individually playing a card game or in a group playing a ball game. Pilot study.	One session/ day for 5 days	Frequency of activity in video observation	1. Improvement in game performance
	(25), (53)	14 subjects (4 men) mean age 79.2, with dementia	Clinic, Japan	Free group interaction with Paro. Pilot study.	One session (20 mins)	EEG recording, Questionnaire	1. Improvement in cortical neurons activity of 7 patients, especially in patients who liked the robot
Social Facilitator	(30)	18 subjects (all female) with dementia	Nursing Home, participants room, US	Supervised one-on-one interaction with AIBO, Dog or no object. Crossover study.	One visit (~ 3 mins)/ week for 3 weeks (each week is a different interaction)	Observed behaviour seen in video-recording	1. All visits generate interactive behaviour with visitor
	(31)	7 subjects with dementia	Dementia rehabilitation wing, US	Supervised group interaction with Paro. Pilot study.	One session (30-45 mins)/ week for 7 weeks	Observed behaviour of primary and non-primary interactor seen in video-recording	1. PARO increases activity in particular modalities of social interaction, which vary between primary and non-primary interactors 2. PARO improved activity levels
	(28)	12 subjects (9 men), mean age 77.25	Residential Care facility, Taiwan	Supervised group interaction with Paro. Pilot study.	Two sessions (30 mins)/ week for 4 weeks	ACIS, Activity Participation scale	1. Significant improvement in communication and interaction skills 2. Significant improvement in activity participation
	(27), (54)	23 subjects, aged 60-104, with high functioning in one nursing home and schizophrenia and/ or dementia in the other	Nursing Homes, US	Supervised group interaction with Paro switched on, Paro switched off or no object. Crossover study.	One session (20 mins)/ 2 weeks (in site A) or per month (in site B) for 4 months (8 sessions vs 4 sessions)	Questionnaire and Observation	1. In switched on Paro group, there was an increase in social interactions; even more in the presence of caregivers or experimenters 2. Switched on Paro also generated feel-good experiences
	(29)	8 subjects (2 men) aged 68-89, with dementia	Group home, Japan	One-on-one interaction with AIBO. Pilot study.	1 session (30 mins)	N-dementia scale, MMSE, behaviour scale and video observation	1. Improving communication with staff in a group home and establishment of friendly relations with occupants
Companionship	(32)	38 subjects	Nursing Home, US	Free one-on-one interaction with AIBO/ dog or no object. Randomised controlled trial.	One session (30 mins)/ week for 8 weeks	Modified LAPS, UCLA LS	1. Dog and AIBO therapy equally reduced loneliness compared to control (more improvement in most lonely participants; in the control group, the most lonely became more lonely) 2. Residents became significantly and equally attached to AIBO and dog. 3. Attachment was not the mechanism for reduced loneliness in dog or AIBO therapy
	(33)	34 subjects, aged >55 years	Retirement Home, New Zealand	Group or individual interaction with Paro or alternative activity. Randomised	Twice (1 hour)/ week for 12 weeks	UCLA LS, GDS, QoLAD	1. Loneliness scores significantly decrease in the Paro group compared to control

Use of Assistive Robot Technology in Elderly Care

				controlled trial.			
	(7)	6 subjects (1 man) aged >64 5 separate control subjects used for CgA measurement	Nursing home/ Participant's home, Japan	Free interaction with AIBO. Control group for CgA measurements had no intervention. Pilot study.	Four sessions (1 hour)/ week for 7 weeks	Scores of emotional words, Amount of speech and Satisfaction, AOKLS, SF-36 and salivary CgA	1. Significant reduction of loneliness 2. Improvement in health related quality of life 3. Decrease in salivary CgA, an indicator of sympathetic adrenal system activity 4. Increase in emotional words, amount of speech and satisfaction exhibited
Physiological Therapy	(34)	21 subjects (7 men) mean age 84.9	Residential care facility, New Zealand	Supervised one-on-one interaction with Paro. Pilot study.	One session (10 mins)	Blood pressure reading: Before during and after interaction	1. Significant reductions in systolic and diastolic blood pressure 2. Reduced systolic blood pressure was sustained after Paro was taken away 3. Reduced diastolic blood pressure was not sustained after Paro was taken away 4. Data suggests average heart rate decreased
	(35), (55), (56), (57), (58), (59)	12 subjects, aged 67-89, with mixed cognitive function	Residential care facility, Japan	Free individual/ group interaction with Paro. Pilot study.	Everyday (9.5 hours) for 4 weeks	Urinary tests, interviews and video recording observation	1. Increase in social interaction and density of social networks 2. Improvement of subjects' vital organs reaction to stress

ABMI, Agitated Behaviours Mapping Instrument; ACIS, Assessment of Communication and Interaction Skills; AES, Apathy Evaluation Scale; AI, Apathy Inventory; AIBO, Artificial Intelligence Robot; AOKLS, Ando Osada and Kodama Loneliness Scale; APADEM-NH, Apathy Scale for Institutionalized Patients with Dementia Nursing Home version; APG, Accelerated Plethysmography; BARS, Brief Agitation Rating Scale; BMI, Body Mass Index; CAM, Confusion Assessment Method; CgA, Chromogranin A; CMAI, Cohen Mansfield Agitation Inventory; Coop/ Wonca, Mood scale; CSDD, Cornell Scale for Symptoms of Depression in Dementia; GBS, Gottfries-Bråne-Steen scale; GDS, Geriatric Depression Scale; GLDS, Global Deterioration Scale; IPPA, Goal attainment scale; LAPS, Lexington Attachment to Pets Scale; LMBS, Lawton's Modified Behaviour Stream; MMSE, Mini Mental State Examination; NAO, NPI, Neuropsychiatric Inventory; OERS, Observed Emotion Rating Scale; POMS, Profile of Mood States; QoLAD, Quality of Life in Alzheimer's Disease Scale; QUALID, Quality of Life Scale; RAID, Rating Anxiety in Dementia Scale; SF-36, Short Form Health Survey; sMMSE, Severe Mini Mental State Examination; TAU, Treatment As Usual; UCLA LS, University of California Los Angeles Loneliness Scale

Table 2: Description of Socially Assistive Robots used in Included Studies

Robot	Description	Number Used In Respective Roles					
		Affective Therapy	Cognitive Training	Social Facilitation	Companionship	Stress Therapy	Total
AIBO	A non-verbal, dog-like robot with a metallic appearance and the ability of sight, walking and interpreting commands. AIBO can learn, mature and, on human interaction, express emotional responses.	-	1	2	2	-	4
Bandit	A humanoid robot mounted on a wheeled base. Bandit can speak, gesticulate and make facial expressions.	-	1	-	-	-	1
JustoCat	A non-verbal, cat-like robot with replaceable fur and similar proportions and weight to a real cat. JustoCat is capable of breathing, purring and meowing and is designed to sit on a persons lap and respond to stroking.	1	-	-	-	-	1
Mero	A humanoid head mounted on a base, capable of head motion, facial expressions and speech.	-	1	-	-	-	1
NAO	A humanoid robot, 58 cm tall, capable of walking, speech, gesticulation and dance. NAO is able to interact with people and can develop new skills and become personalised.	1	1	-	-	-	2
NeCoRo	A non-verbal, cat-like robot designed to move and look like a real cat. NecoRo can interpret its surroundings and move accordingly. NeCoRo can express emotion.	1	-	-	-	-	1
Nodding Kabochan	A small robot, with the appearance of a child-like teddy, that can talk, sing and nod. It is designed to communicate with users. Nodding Kabochan can play exercise and singing games with the user.	-	1	-	-	-	1
Silbot	A penguin-like robot that can speak and detect faces. Silbot can engage with users in conversation, games and provide care through drug regimen reminders.	-	1	-	-	-	1
Paro	A non-verbal, seal-like robot with the ability to move its head and tail, blink and make sounds and has 5 sensory modalities: light, sound, temperature, posture and tactile. Paro will respond to being held or stroked and can learn to respond to its name. Paro has its own rhythms; will at times be playful, and at other times sleepy and inactive.	9	2	3	1	2	17

Use of Assistive Robot Technology in Elderly Patients

Table 3: Data extracted from comparable studies in Affective Therapy studies

AFFECTIVE THERAPY Mood Scores									
			Control			**Intervention**			
Study	Number of subjects	Outcome Scale	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	p-value
(17)	4	CMAI	-	-	-	12.6 (6.3)	13.3 (6.6)	0.7	0.88[a]
(13)	53	BARS	22 (19)	23.3 (22)	1.3	20.1 (12.8)	13.7 (11.7)	-6.4	0.044[b]
(13)	53	CSDD	18.2 (12.3)	24.5 (17.3)	6.3	23.7 (12.9)	18.9 (16.8)	-4.8	0.019[b]
(12)	18	GDS	-	28.7 (23.3)	-	-	31.3 (19.3)	-	0.72[c]
(19)	100	GDS	-	-	-	13.3 (6.7; 33.3)*	13.3 (6.7; 23.3)*	-	< 0.05[d]

Table 4: Data extracted from comparable studies in Cognitive Training studies

COGNITIVE TRAINING Cognition Scores									
			Control			**Intervention**			
Study	Number of subjects	Outcome Scale	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	p-value
(22)	34	MMSE	-	-	-	94.0 (5)	99.0 (2.3)	5	< 0.01[a]
(21)	101	MMSE	12.1 (18.1)	10.4 (15.7)	-1.7	11.8 (17.3)	8.1 (15.0)	-3.7	0.022[b]
(21)	101	MMSE	12.1 (18.1)	10.4 (15.7)	-1.7	10.7 (16.5)	9.1 (15.7)	-1.6	0.282[b]
(24)	71	ADAS-Cog	-	-	-	89.9 (5.1)	92.6 (4.0)	2.7	<0.001[a]

Use of Assistive Robot Technology in Elderly Patients

Table 5: Data extracted from comparable studies in Companionship studies

COMPANIONSHIP Loneliness Scores									
			Control			**Intervention**			
Study	Number of subjects	Outcome Scale	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score (SD)	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score (SD)	p-value
(32)	38	UCLA LS	-	-	5.7 (1.3)	-	-	-6.0 (2.7)	< 0.05[b]
(33)	34	UCLA LS	-	-	3.8 (10.3)	-	-	-9.0 (12.6)	0.03[b]
(7)	5	AOKLS	-	-	-	3.3 (2.2)	1.0 (1.3)	-	0.07[a]

Use of Assistive Robot Technology in Elderly Patients

a – Study compares mean baseline score in intervention group to mean follow-up score in the intervention group

b – Study compares change in mean score from baseline to follow-up in control group to change in mean score from baseline to follow-up in intervention group

c – Study compares mean follow up score of control group to mean follow up score of intervention group

d – Study compares median baseline score in intervention group to median follow-up score in the intervention group

* Median and interquartile range reported

ADAS-Cog, Alzheimer's Disease Assessment Scale – Cognitive subscale; AOKLS, Ando Osada and Kodama Loneliness Scale; BARS, Brief Agitation Rating Scale; BMI, Body Mass Index; CMAI, Cohen Mansfield Agitation Inventory; CSDD, Cornell Scale for Symptoms of Depression in Dementia; GDS, Geriatric Depression Scale; MMSE, Mini Mental State Examination; UCLA LS, University of California Los Angeles Loneliness Scale.

Discussion

The aim of this review is to identify the roles SAR could play in elderly care. Despite the infancy of this field, the qualitative amalgamation of the studies demonstrated 5 roles for SAR, notably as Affective Therapy and Cognitive Training tools.

Evaluation of SAR Technology

This review provides the first holistic insight into what roles SAR can play in elderly care with the identification of 5 uses. Notwithstanding promising results, the evidence base behind each role varies substantially with Affective Therapy and Cognitive Training roles of SAR having the strongest evidence base. However, it would be unfair to dismiss the remaining 3 sets of data given the novelty of the field. Although this review did not exclude studies based on methodology, there is much this data offers to highlight the technological potential of SAR in elderly care. For example, a Companionship role is popular concept for SAR among commentators in the literature, but very few actual studies have been conducted to demonstrate this. The evidence for this role is desirable and the multitude of applications for a companion robot is one of many reasons why this concept has gained popularity. As reported by one of the selected studies(32), AIBO, the robotic dog, was as effective a companion as a real dog. This has real implications for its use, specifically where a real animal companion may be inappropriate.

Similarly, evidence for the Social Facilitator role was only reported in smaller, methodologically weaker studies. However, the potential utility of such a role is substantial. A major concern of SAR is the replacement of human interaction. The limited evidence thus far suggests that the robots may instead be able to increase social activity amongst the elderly. At present, it is difficult to establish precisely how SAR can improve the sociability of

elderly users, and indeed the excitement alone of a new robot in the nursing home could be enough to stimulate new conversations. Larger studies, using validated outcome measures, need to be conducted to clarify how SAR improves the users' sociability.

Affective Therapy and Cognitive Training have perhaps the most promise for future SAR use, with the potential to improve the general wellbeing and independence of elderly users. Therapeutic tools to combat mood disturbances or maintain cognitive function through old age could prove to be very useful. From a conversationalist partner in the hospital to a group teacher in the nursing home, SAR could re-imagine what healthcare looks like from a senior's perspective. The current weight of evidence suggests that SAR can improve mood and cognitive function in the short term, which could have applications in acute periods of distress or to supplement short term care in a hospital setting.

Ethical Considerations

It would be prudent to consider the ethical facets of such a novel form of intervention because whilst SAR may have demonstrated utility in certain roles, it does not necessarily mean it should fulfil them. For example, consider the Companionship role. Assuming SAR were able to provide long term protection against loneliness and isolation, it risks reducing the actual amount of contact the user has with other people. Some have argued that simply having the robot to talk to could be a justification for relatives and friends to postpone visiting(60). There are also the concerns associated with presenting a robot as though it can form a real relationship. One report(61) highlighted that the entire intervention is predicated on deceiving the user into thinking the robot is a sentient being. Such behaviour raises moral questions.

These concerns however, tend not to reflect the current evidence of SAR use, but rather are based on the expectations of what the technology may offer in the future. For example, one of the roles identified was the Social Facilitator role. Certainly if SAR interventions can increase sociability and human interaction, the concern of inadvertent isolation may be irrelevant. It is nonetheless important to ensure that any application of SAR does not come at the expense of reduced human interaction.

Furthermore, the value of a relationship between a person and a robot may not inherently be a bad one, and does not have to be predicated on illusion. As pointed out in one report (60), it is very possible for an individual to enjoy the company and interaction of a robot without necessarily believing it is sentient. Given the very limited capabilities of the existing robots, much of the ethical concern tends to resemble a reality more akin to science fiction, and is broadly unsubstantiated in the current evidence. Nonetheless, as the field develops, one expects so too will the concerns.

Limitations of the Selected Studies

Whilst positive results have been reported, there are several underlying methodological limitations facing the selected studies. This complicates the task of establishing the clinical application of such technologies, and risks undermining the field's efforts or sensationalising exploratory research. The first limitation is that the use of SAR in elderly care is still a small field of study and despite using wide search criteria, only a handful of relevant studies were found. Indeed, many of the studies identified were reported in conference proceedings and utilised a narrow set of robots, primarily Paro and AIBO. This dramatically restricts the quantity and quality of information available to validate or dismiss much of the futuristic speculation that surrounds SAR.

There is also the concern of cultural bias since a large proportion of the studies were conducted in Japan. Although more recent studies have been conducted in other cultural environments, most notably the US, they too have methodological challenges in their study design, with many having small sample sizes, no control, no randomisation and very short study durations. Furthermore, there is evidence of gender bias, with a disproportionate number of women included among many of the selected studies. This is a concern since men and women as populations tend to regard robot technology differently(62), and as such some of the reported findings may be exaggerated or diminished by the participant composition.

Another, perhaps subtle, issue that is found in 15 of the studies is the supervision of interactions. Although supervision ensures safety, it risks altering how the participant interacts with the robot and may change how the participant reports the robot's utility; known as the Hawthorn Effect. Whilst this is difficult to control for when the study is not randomised and no comparator is used, direct supervision may lead to greater positive effects reported than is necessarily the case.

Similarly, the utility of the reported data is further hampered by the short-term demonstration of the reported outcomes and it is still not clear whether these results translate into long term improvements. Whilst there could be a demand for short term interventions in certain settings (hospital stays or overcoming acute periods of mood disturbance), the absence of long term data restricts the practicality of introducing expensive SAR into mainstream elderly care.

These challenges are further compounded by the nature of the studies' outcome measures. These are often abstract, with a limited number of studies identifying a direct clinical need or problem, instead focussing on desired applications and outcomes. In many

of the cases where a control comparator was used, it often involved uninspiring activities or no activity at all which is an unfair comparison, and may inflate the value attributed to the results. As momentum grows behind SAR, these concerns will need to be addressed if the technology is going to play a significant clinical role in the future.

Review Limitations

The primary limitation of this review is the validity of the categorisation of studies into the defined roles. The roles were created retrospectively, as part of a discovery process on extracting data from the final set of studies. Whilst they have utility in evaluating the state of the field and providing defined expectations for the technology, they have generalised sets of studies that are very different in quality, design and sometimes outcome. There is also the issue that some studies demonstrated several roles for SAR, however have been limited and categorised into the role which best matches the study's primary outcome measure. As a consequence, this may mislead the actual weight of data in the respective roles, despite the transparency of the reported outcomes in Table 1.

Additionally, this review ran the inadvertent risk of excluding relevant papers in the screening phase. Although high concordance between the reviewers was reported, the large volume of studies that had to be reviewed invites the possibility that relevant publications were excluded. The main reason for the high exclusion rate was because the broad search criteria identified irrelevant robot interventions, such as surgical robots or telecommunication devices. It is unlikely, however, that an additional study would have changed the conclusions the review came to.

Finally, the comparison of assessment values between studies illustrated in Tables 3-5, whilst providing some insight into the quality of evidence in the field, is not a meaningful

comparison, because of the varying study designs and scales used. This has made it very difficult to come to any substantial conclusions about the identified roles.

Future of the Field

In order to achieve successful application of SAR in elderly care, future studies should be more conscious of the outcome measure chosen and its translation into care. Some studies used surrogate measures such as frequency of laughter(18), or performance in particular games(26). Whilst these may be desired outcomes, it is not clearly demonstrated how they meet quantifiable needs of the elderly population. It is likely that any application of SAR will incorporate several of the previously defined roles, therefore larger studies should assess the intervention's impact in the context of these clear roles with validated outcome measures. For example, one study(22) involved a robot staying at home with the elderly participants for 8 weeks, and assessed its impact using questionnaires, cognitive tests, blood and saliva samples and an accelerated plethysmography. Whilst the study demonstrated an improvement in cognitive scores and a reduction in saliva cortisol, it did not assess whether living with a robot for 8 weeks had any impact on loneliness. This leaves the companionship capabilities of this type of SAR intervention unclear. Larger randomised controlled trials using valid comparators are needed to definitively show where SAR is and is not useful in elderly care.

Conclusion

Socially assistive robots have shown potential in elderly care which, in light of an unprecedented demographic shift, promises to reform the delivery of care for the elderly. Although many of the studies described carry their own methodological issues, the size and

quality of studies are improving. This review has qualitatively assessed the existing research and comprehensively outlined the state of the field as it stands. In establishing the 5 roles to which SAR can be ascribed, this review intends not to restrict ambition, but to provide a basis for clinical applicability and design of future studies. This review urges that new studies should be clearer about the precise role any robot intervention intends to serve, and use validated measures to assess their effectiveness. Future studies need to demonstrate how SAR can solve real problems in order to shift from novelty to functionality in elderly care.

Acknowledgements

This review would like to thank Peter Forrest for his scientific guidance and participation in discussions for this project, and Sundhiya Mandalia for her support in interpreting of the data and its analysis.

Declaration of competing interests: None

Contribution (Author contributing):

Design of the review (Jordan Abdi, Marcela Vizcaychipi)
Data collection (Jordan Abdi, Ahmed Al-Hindawi)
Data analysis and interpretation (Jordan Abdi, Tiffany Ng)
Drafting the article (Jordan Abdi, Ahmed Al-Hindawi, Tiffany Ng)
Critical revision of the article (Jordan Abdi, Tiffany Ng, Marcela Vizcaychipi)
Final approval of the version to be published (Marcela Vizcaychipi)

Funding sources: All authors declare independence from funders.

Transparency Declaration: All authors, external and internal, had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. The lead author affirms the manuscript is honest, accurate and that any discrepancies from the study are accounted for.

Data sharing: no additional data available.

Systematic registration number and name of register: NIHR 58672

References

1. United Nations. World Population Prospects. 2015;1. Available from: http://esa.un.org/unpd/wpp/Publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf
2. Thomas J. Insights into Loneliness , Older People and Well-being. 2015;1–10.
3. Victor C, Scambler S, Bond J, Bowling A. Loneliness, social isolation and living alone in later life. Sheffield; 2003;
4. Holt-lunstad J, Smith TB, Layton JB. Social Relationships and Mortality Risk : A Meta-analytic Review. 2010;7(7).
5. Hegel F, Muhl C, Wrede B, Hielscher-Fastabend M, Sagerer G. Understanding Social Robots. Second Int Conf Adv Comput Interact. 2009;(Section II):169--174.
6. Reiser U, Jacobs T, Arbeiter G, Parlitz C. Care-O-bot ® 3 – Vision of a Robot Butler. 2013;97–116.
7. Kanamori M, Suzuki M, Oshiro H. Pilot study on improvement of quality of life among elderly using a pet-type robot. 2003;107–12.
8. Wada K, Shibata T, Saito T, Tanie K. Effects of Robot Assisted Activity to Elderly People who Stay at a Health Service Facility for the Aged. 2003;(October).
9. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. 2011. Available from: http://handbook.cochrane.org
10. Orwin RG, Vevea JL. Evaluating Coding Decisions. In: The Handbook of Research Synthesis and Meta-Analysis. 1994. p. 177–200.
11. Moher D, Liberati A, Tetzlaff J, Altman DG, Grp P. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement (Reprinted from Annals of Internal Medicine). PLoS Med [Internet]. 2009;89(9):873–80. Available from: <Go to

Use of Assistive Robot Technology in Elderly Patients

ISI>://000270260200002

12. Moyle W, Cooke M, Beattie E, Jones C, Klein B, Cook G, et al. Exploring the effect of companion robots on emotional expression in older adults with dementia: a pilot randomized controlled trial. J Gerontol Nurs [Internet]. United States; 2013;39(5):46–53. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=23506125

13. Joranson N, Pedersen I, Rokstad AMM, Ihlebaek C. Effects on Symptoms of Agitation and Depression in Persons With Dementia Participating in Robot-Assisted Activity: A Cluster-Randomized Controlled Trial. J Am Med Dir Assoc [Internet]. United States; 2015;16(10):867–73. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=26096582

14. Bemelmans R, Gelderblom GJ, Jonker P, de Witte L. Effectiveness of Robot Paro in Intramural Psychogeriatric Care: A Multicenter Quasi-Experimental Study. J Am Med Dir Assoc [Internet]. United States; 2015;16(11):946–50. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=26115817

15. Wada K, Shibata T, Saito T, Tanie K. Effects of robot assisted activity for elderly people at day service center and analysis of its factors. Proc 4th World Congr Intell Control Autom (Cat No02EX527) [Internet]. 2002;2:1301–5. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1020792

16. Wada K, Shibata T, Saito T, Sakamoto K, Tanie K. Psychological and Social Effects of One Year Robot Assisted Activity on Elderly People at a Health Service Facility for the Aged. Proc IEEE Int Conf Robot Autom. 2003;(April):2785–90.

17. Gustafsson C, Svanberg C, Müllersdorf M. Using a Robotic Cat in Dementia Care. J Gerontol Nurs [Internet]. Senior Lecturer, School of Health, Care, and Social Welfare, Mälardalen University, Eskilstuna and Västerås, Sweden: SLACK Incorporated; 2015 Oct;41(10):46–56 11p. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=110074731&site=ehost-live

18. Takayanagi K, Kirita T, Shibata T. Comparison of Verbal and Emotional Responses of Elderly People with Mild/Moderate Dementia and Those with Severe Dementia in Responses to Seal Robot, PARO. Front Aging Neurosci [Internet]. Switzerland; 2014;6:257. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=25309434

19. Thodberg K, Sorensen LU, Christensen JW, Poulsen PH, Houbak B, Damgaard V, et al. Therapeutic effects of dog visits in nursing homes for the elderly. [Internet]. Psychogeriatrics. United Kingdom: Wiley-Blackwell Publishing Ltd.; 2015. p. No-Specified. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc11&NEWS=N&AN=2015-49600-001

20. Libin A, Cohen-Mansfield J. Therapeutic robocat for nursing home residents with dementia: preliminary inquiry. Am J Alzheimers Dis Other Demen [Internet]. United States; 2004;19(2):111–6. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med5&NEWS=N&AN=15106392

21. Valentí Soler M, Agüera-Ortiz L, Olazarán Rodríguez J, Mendoza Rebolledo C, Pérez Muñoz A, Rodríguez Pérez I, et al. Social robots in advanced dementia. Front Aging Neurosci [Internet]. Frontiers; 2015 Jan 3 [cited 2016 Mar 14];7:133. Available from:

http://journal.frontiersin.org/article/10.3389/fnagi.2015.00133/abstract

22. Tanaka M, Ishii A, Yamano E, Ogikubo H, Okazaki M, Kamimura K, et al. Effect of a human-type communication robot on cognitive function in elderly women living alone. Med Sci Monit [Internet]. Poland; 2012;18(9):CR550-7. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=22936190

23. Tapus A. Improving the quality of life of people with dementia through the use of socially assistive robots. Proc - 2009 Adv Technol Enhanc Qual Life, AT-EQUAL 2009. 2009;81–6.

24. Kim GH, Jeon S, Im K, Seo SW, Cho H, Noh Y, et al. Structural brain changes after robot-assisted cognitive training in the elderly: A single-blind randomized controlled trial. Alzheimer's Dement [Internet]. 2013 Jul [cited 2016 Mar 14];9(4):P476–7. Available from: http://www.sciencedirect.com/science/article/pii/S1552526013016282

25. Wada K, Shibata T, Musha T, Kimura S. Robot Therapy for Elders Affected by Dementia. 2014 IEEE Long Isl Syst Appl Technol Conf LISAT 2014. 2014;(August).

26. Hamada T, Okubo H, Inoue K, Maruyama J, Onari H, Kagawa Y, et al. Robot therapy as for recreation for elderly people with dementia - Game recreation using a pet-type robot -. Proc 17th IEEE Int Symp Robot Hum Interact Commun RO-MAN. 2008;174–9.

27. Kidd CD, Taggart W, Turkle S. A Sociable Robot to Encourage Social Interaction among the Elderly.

28. Sung H-C, Chang S-M, Chin M-Y, Lee W-L. Robot-assisted therapy for improving social interactions and activity participation among institutionalized older adults: a pilot study. Asia Pac Psychiatry [Internet]. Australia; 2015;7(1):1–6. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24692085

29. Sakairi K. Research of robot-assisted activity for the elderly with senile dementia in a group home. SICE 2004 Annu Conf. 2004;3:2092–4.

30. Kramer SC, Friedmann E, Bernstein PL. Comparison of the effect of human interaction, animal-assisted therapy, and AIBO-assisted therapy on long-term care residents with dementia. [Internet]. Banks Beck, Bernstein, Churchill, Cole, Crowley-Robinson, Fick, Fritz, Kongable, Lehner, Lutwack-Bloom, Martin, McCabe, Melson, Motomura, Perelle, Richeson, Smith, Stoppe B, editor. Anthrozoos. Kramer, Stephen C., 58 Jackson Road, Bedford, NY, US, 10506, Stephenkramer5@aol.com: Berg Publishers; 2009. p. 43–57. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc6&NEWS=N&AN=2009-04201-004

31. Sabanovic S, Bennett CC, Chang W-L, Huber L. PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. IEEE Int Conf Rehabil Robot [Internet]. United States; 2013;2013:6650427. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24187245

32. Banks MR, Willoughby LM, Banks WA. Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. J Am Med Dir Assoc [Internet]. United States; 2008;9(3):173–7. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med5&NEWS=N&AN=18294600

33. Robinson H, Macdonald B, Kerse N, Broadbent E. The psychosocial effects of a companion robot: a randomized controlled trial. J Am Med Dir Assoc [Internet]. United States; 2013;14(9):661–7. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=23545466

34. Robinson H, MacDonald B, Broadbent E. Physiological effects of a companion robot on blood

pressure of older people in residential care facility: a pilot study. Australas J Ageing [Internet]. Australia; 2015;34(1):27–32. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24373064

35. Wada K, Shibata T. Social and physiological influences of robot therapy in a care house. [Internet]. Baum Folstein, Fujita, Gammonley, Kanamori, Libin, Nishikaze, Ohokubo, Robins, Saito, Selye, Shibata, Shibata, Shibata, Shibata, Shibata, Shibata, Shibata, Shibata, Shibata, Shibata, Shibata, Shibata, Tamura, Wada, Wada, Wada, Wada, Werry, Yokoyama, Zunz B, editor. Special Issue: Human and robot interactive communication. Wada, Kazuyoshi: Intelligent Systems Research Institute, AIST, 1-1-1 Umezono, Tsukuba, Ibaraki, Japan, 305-8568, k-wada@aist.go.jp: John Benjamins; 2008. p. 258–76. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc6&NEWS=N&AN=2008-08717-006

36. Wada KWK, Shibata TST, Saito TST, Sakamoto KSK, Tanie KTK. Robot assisted activity at a health service facility for the aged for 17 months: an interim report of long-term experiment. IEEE Work Adv Robot its Soc Impacts, 2005. 2005;127–32.

37. Wada K, Shibata T, Sakamoto K, Tanie K. Quantitative analysis of utterance of elderly people in long-term robot assisted activity. Proc - IEEE Int Work Robot Hum Interact Commun. 2005;2005:267–72.

38. Wada K, Shibata T, Saito T, Tanie K. Psychological and social effects in long-term experiment of robot assisted activity to elderly people at a health service facility for the aged. 2004 IEEE/RSJ Int Conf Intell Robot Syst [Internet]. 2004;3:3068–73. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-14044273830&partnerID=40&md5=b85e82f0c643a517d76b6eb7a05fe686

39. Wada K, Shibata T, Kawaguchi Y. Long-term robot therapy in a health service facility for the

aged-A case study for 5 years. 2009 IEEE Int Conf Rehabil Robot ICORR 2009 [Internet]. 2009;930–3. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-70449380117&partnerID=40&md5=4bba7ff889ec97a3bd7ebd52ab8af615\nhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5209495\nhttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5209495

40. Wada K, Shibata T, Sakamoto K, Saito T, Tanie K. Analysis of utterance in long-term robot assisted activity for elderly people. Proceedings, 2005 IEEE/ASME Int Conf Adv IntelligentWada, K, Shibata, T, Sakamoto, K, Saito, T, Tanie, K (2005) Anal utterance long-term Robot Assist Act Elder people Proceedings, 2005 IEEE/ASME Inte [Internet]. 2005;1(April):31–6. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1500961

41. Wada K, Shibata T, Saito T, Tanie K. Robot assisted activity at a health service facility for the aged for ten weeks: an interim report of a long-term experiment. Int IEEE/EMBS Conf Neural Eng NER. 2006;470–3.

42. Wada K, Shibata T, Saito T, Tanie K. Effects of Three Months Robot Assisted Activity to Depression of Elderly People who stay at a Health Service Facility for the Aged. SICE 2004 Annu Conf. 2004;2709–14.

43. Wada K, Shibata T, Saito T, Tanie K. Analysis of factors that bring mental effects to elderly people in robot assisted activity. IEEE/RSJ Int Conf Intell Robot Syst. 2002;2(October):1152–7.

44. Wada K, Shibata T, Saito T, Tanie K. Robot Assisted Activity for Elderly People and Nurses at a Day Service Center. IEEE/RSJ. 2002;(May):1416–21.

45. Wada K, Shibata T, Saito T, Tanie K. Robot assisted activity to elderly at a health service facility for the aged. Int IEEE/EMBS Conf Neural Eng NER. 2003;2003–Janua:470–3.

46. Saito T, Shibata T, Wada K, Tanie K. Relationship between interaction with the mental commit

robot and change of stress reaction of the elderly. Proc IEEE Int Symp Comput Intell Robot Autom CIRA. 2003;1:119–24.

47. Wada K, Shibata T, Saito T, Tanie K. Relationship between familiarity with mental commit robot and psychological effects to elderly people by robot assisted activity. Proc IEEE Int Symp Comput Intell Robot Autom CIRA. 2003;1:113–8.

48. Wada K, Shibata T, Saito T, Tanie K. Psychological and Social Effects of Robot Assisted Activity to Elderly People who stay at a Health Service Facility for the Aged. IEEE Int Conf Robot Autom. 2004;3:14–9.

49. Wada K, Shibata T, Saito T, Tanie K. Effects of robot-assisted activity for elderly people and nurses at a day service center. Proc IEEE. 2004;92(11):1780–8.

50. Tapus A, Ţăpuş C, Matarić MJ. The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. 2009 IEEE Int Conf Rehabil Robot ICORR 2009. 2009;924–9.

51. Tapus A, Tapus C, Mataric M. Music therapist robot for individuals with cognitive impairments. 2009;297–8.

52. Tapus A, Tapus C, Mataric M. Music Therapist Robot: A Solution for Helping People with Cognitive Impairments. 2009;297–8.

53. Wada K, Shibatal T, Musha T, Kimura S. Effects of robot therapy for demented patients evaluated by EEG. 2005 IEEE/RSJ Int Conf Intell Robot Syst IROS. 2005;2205–10.

54. Taggart W, Turkle S, Kidd CD. An Interactive Robot in a Nursing Home : Preliminary Remarks Field Setting : Nursing Homes. Towar Soc Mech Android Sci A COGSCI Work. 2005;1–6.

55. Wada K, Shibata T. Social effects of robot therapy in a care house - Change of social network of the residents for two months. Proc - IEEE Int Conf Robot Autom. 2007;(April):1250–5.

56. Wada K. Social an physiological influences of living with seal robots in an elderly care house for

two months. Gerontechnology [Internet]. 2008;(June 2005):235–235. Available from: http://www.gerontechnology.info/Journal/pdf.php?a=645

57. Wada K, Shibata T. Robot therapy in a care house - Change of relationship among the residents and seal robot during a 2-month long study. Proc - IEEE Int Work Robot Hum Interact Commun. 2007;107–12.

58. Wada K, Shibata T. Robot therapy in a care house - Results of case studies. Proc - IEEE Int Work Robot Hum Interact Commun. 2006;(June 2005):581–6.

59. Wada K, Shibata T. Living With Seal Robots—Its Sociopsychological and Physiological Influences on the Elderly at a Care House. Proc - IEEE Int Work Robot Hum Interact Commun. 2007;23(5):972–80.

60. Sharkey A, Sharkey N. Granny and the robots: Ethical issues in robot care for the elderly. Ethics Inf Technol. 2012;14(1):27–40.

61. Sparrow R. The march of the robot dogs. Ethics Inf Technol [Internet]. 2002;305–18. Available from: http://link.springer.com/article/10.1023/A:1021386708994

62. Schermerhorn P, Scheutz M, Crowell CR. Robot social presence and gender: Do females view robots differently than males? 3rd ACM/IEEE Int Conf. 2008;

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	11
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	11
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	11
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	19
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	10
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	19
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	29
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	31
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	34
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	36

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

BMJ Open

A Systematic Review on the Use of Socially Assistive Robot Technology in Elderly Care

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-018815.R1
Article Type:	Research
Date Submitted by the Author:	17-Oct-2017
Complete List of Authors:	Abdi, Jordan; Imperial College London, Al-Hindawi, Ahmed; Chelsea and Westminster Hospital, Magill Department of Anaesthesia Ng, Tiffany; Chelsea and Westminster Hospital, Magill Department of Anaesthesia Vizcaychipi, M; Imperial College London, Magill Deaprtment Anaesthesia
Primary Subject Heading:	Geriatric medicine
Secondary Subject Heading:	Sociology, Public health, Mental health
Keywords:	Socially Assistive Robots, Elderly, Robots, GERIATRIC MEDICINE, SOCIAL MEDICINE

A Systematic Review on the Use of Socially Assistive Robot Technology in Elderly Care

Jordan Abdi[1], Ahmed Al-Hindawi[2], Tiffany Ng[2], Marcela P Vizcaychipi[2]

[1]Department of Medicine, Imperial College London, Kensington, London SW7 2AZ

[2]Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH

Author note:

Jordan Abdi (Abdi J), Undergraduate student, Department of Medicine, Imperial College London

Ahmed Al-Hindawi (Al-Hindawi A), Intensive Care Registrar, Magill Research Fellow, Magill Department of Anaesthesia, Chelsea and Westminster Hospital

Tiffany Ng (Ng T), Doctor, Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH.

Marcela Vizcaychipi (Vizcaychipi MP), Consultant in Anaesthesia and Intensive Care in the Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH

Address Correspondence: Jordan Abdi, Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH.

Electronic address: ja1612@ic.ac.uk

Use of Assistive Robot Technology in Elderly Care

STRUCTURED ABSTRACT

OBJECTIVE: With an elderly population that is set to more than double by 2050 worldwide, there will be an increased demand for elderly care. This trend is paired with a fall in the number of younger people able to support the older members of society. This poses several impediments in the delivery of high quality health and social care. Socially Assistive Robot (SAR) technology could assume new roles in health and social care to meet this higher demand. This review qualitatively examines the literature on the use of SAR in elderly care and aims to establish the roles this technology may play in the future.

DESIGN: Systematic Review

DATA SOURCES: Systematic search of CINAHL, Cochrane Library, EMBASE, MEDLINE, PsychINFO and SCOPUS databases was conducted, complemented with a free search using Google Scholar and reference harvesting. All publications went through a selection process, based on pre-defined selection criteria, which involved sequentially reviewing the title, abstract and full text of the publication. No limitations regarding date of publication were imposed and only English publications were taken into account.

ELIGIBILITY CRITERIA: The inclusion criteria consist of elderly participants, any elderly healthcare facility, humanoid and pet robots, and all social interaction types with the robot. Exclusions were acceptability studies, technical reports of robots and publications surrounding physically or surgically assistive robots.

RESULTS: In total, 61 final publications were included in the review, describing 33 studies and including 1574 participants and 11 robots. 28 of the 33 papers report positive findings. Five roles of SAR were identified: Affective Therapy, Cognitive Training, Social Facilitator, Companionship and Physiological Therapy.

CONCLUSIONS: Although many positive outcomes were reported, a large proportion of the studies have methodological issues, which limit the utility of the results. Nonetheless, the reported value of SAR in elderly care does warrant further investigation. Future studies should endeavour to validate the roles demonstrated in this review.

SYSTEMATIC REVIEW REGISTRATION - NIHR 58672

Abstract Word Count: 312 Words

Keywords: Socially assistive robots, elderly, robots, geriatric care, social care

STRENGTHS AND LIMITATIONS

- This is the first systematic review of the literature that has evaluated and categorised the effects of Socially Assistive Robot interventions aimed to improve the health and social care of elderly people.
- The novelty of the field means that the quantity and quality of studies available in the current literature is limited, making generalisations difficult.
- The retrospective creation of SAR roles grouped together sets on studies that differed in quality, design and sometimes outcome, which may mislead the actual weight of data in the respective roles.

INTRODUCTION

The global population is undergoing a demographic shift. Life expectancy is growing and the post-war baby boom generation is entering retirement. The implications on resource allocation will impact the delivery of elderly care. As of 2015(1), 21% of Western Europe's population were over the age of 60, and this is expected to rise to 33% by 2030. By 2050, there are expected to be more people over the age of 60 globally than under 15, reaching a total population of 2.1 billion compared to 901 million in 2015. This is compounded by a proportional decrease in the number of social and health care providers shouldering this increased burden. In 2015, 7 workers were allocated for every elderly person globally, but this is projected to fall to 4.9 in 15-years(1). Moreover, the situation is magnified in Europe by an accelerated ageing population. Currently, there are 3.5 workers for every elderly person, but this is set to fall to 2.4 by 2030. The shift in societal proportions will place new pressures on all aspects of elderly care.

Loneliness, for instance, is a consequence of social, psychological and personal factors. Over half of people over the age of 75 live alone(2) and 17% of older people see family, friends or neighbours less than once a week(3). A recent meta-analysis(4) showed that the impact of loneliness and isolation carries the same mortality risk as smoking 15 cigarettes a day. This is compounded by the fact that social care is a labour intensive industry in a world with a proportionally shrinking workforce.

Throughout many industries the 'robot revolution' promises to solve this growing personnel shortage. At present, physically or surgically assistive robots dominate the healthcare sector's robot usage. This includes: (i) increasingly sophisticated wheelchairs transforming the limitations imposed on paraplegics; (ii) robotic limbs redefining amputee capabilities; and (iii) robotic surgeons revolutionising how and where surgery can be

performed. Nonetheless, physically assistive robots do not combat the increasing mental health burden recognised in the elderly population. It is here that the concept of socially assistive robots is gaining headway. These are robots adept at completing a complex series of physical tasks with the addition of a social interface capable of convincing a user that the robot is a social interaction partner(5).

Socially assistive robots have been categorised into 2 operational groups: 1) service robots and 2) companion robots. Service robots are tasked with aiding activities of daily living(6). Companion robots, by contrast, are more generally associated with improving the psychological status and overall well-being of its users. Such examples include Sony's AIBO(7) and Paro(8). Despite much of the hype, the utilisation of this technology in elderly care is not completely ascertained.

The aim of this systematic review is to establish the clinical usefulness of Socially Assistive Robots (SAR) in elderly care. Through examination and qualitative analyses of existing literature, studies will showcase the utility of SAR and their associated clinical outcomes. A better understanding of SAR and its ability to provide integral care, both socially and physiologically, will provide an indication of its future role in society.

Use of Assistive Robot Technology in Elderly Care

METHODOLOGY

The protocol for this review was conducted in accordance with the principles of the Cochrane Handbook for Systematic Reviews of Interventions (9).

Search Strategy

The following bibliographical databases were searched: CINAHL, Cochrane Library, EMBASE, MEDLINE, PsychINFO and SCOPUS using medical subject headings (MeSH or where appropriate, the database specific thesaurus equivalent) or text word terms. The database search query was composed of 2 search concepts: the intervention (SAR) and the context (elderly care). Free text terms for the intervention included: “service robot*”, “therapeutic robot*” and “socially assistive robot*”; their associated MeSH terms were “Robotics” and “Artificial Intelligence”. The names of specific robot systems were also searched for. The free words used for the context included: “elder*”, “senior*”, “older person*”, “old people” and “dementia”; their associated MeSH term was “Aged, 80 and over”. The use of the asterisk (*) enables the word to be treated as a prefix. For example, “elder*” will represent “elderly” and “eldercare” amongst others. (see supplementary material for an example of a bibliographical search). Additional studies were selected through a free search (Google Scholar) and from reference lists of selected publications and relevant reviews. The main search was conducted in March 2016, and the latest search was conducted in September 2017.

Study Selection

Two reviewers (JA and AA) independently screened the publications in a three-step assessment process: the title, abstract and full text and selection were made in accordance

with inclusion criteria. All publications collected during the database search, free search and reference list harvesting were scored on a 3-point scale (0 = Not relevant, 1 = Possibly relevant, 2 = Very relevant) and those with a combined score of 2 between the reviews would make it through to the next round of scoring. All publications with a total score of 0 were excluded. A publication with a combined score of 1 indicated a disagreement between the reviewers, and would be resolved through discussion. At the end of the full text screening round, a final set of publications to be included into the review was acquired. Cohen's kappa coefficient was calculated to ascertain the agreement between the reviewers in the title, abstract and full text screening phases.

A study was considered eligible if it assessed the usefulness of SAR in the elderly population with a clinical outcome measure. A study that simply assessed the robot's acceptability to elderly users without a clinical outcome measure, or was a technical report, or concerned the use of physically or surgically assistive robots was excluded. No limitations regarding date of publication were imposed and only English publications were considered.

Since the field of Socially Assistive Robotics is in its infancy, many of the studies are small and exploratory. Nonetheless, they provide an insight into what is currently being researched and the potential applications of SAR in elderly care. For this reason, no publication was excluded on the grounds of methodological quality.

Data Extraction

The data extraction form was designed in line with the PICO approach (Participants, Intervention, Comparator and Outcomes). This process was conducted by one reviewer (JA) to ensure consistent extraction of all studies. All clinical outcome measures reported in selected studies were extracted. Data extraction included, in addition to outcomes, country

in which study was conducted, number of included participants, mean age of participants, gender ratio of participants, specific robot used, cognitive status of participants, settings, study design, study duration, and assessment tools.

Duplicate reports of the same study may present in different journals, papers or conference proceedings, and may each focus on different outcome measures or include a follow up data point. To minimise the impact of duplicates, the final set of publications were collated into "study groups" containing duplicate reports. The data extraction process was conducted on the most comprehensive report of a given study.

Data Synthesis and Analysis

Studies were categorised into groups by the role of the robot in the study. The categories were generated retrospectively by the authors and were not pre-defined or directly referenced in the original studies themselves.

Some studies used comparable quantitative outcome measures in their assessment of clinical utility of SAR robots. As different assessment tools were used across studies, a standardised mean score (0 – 100) was generated to allow comparison across similar assessment tools. The result is a unit-free size.

Use of Assistive Robot Technology in Elderly Care

RESULTS

Search Results

The database search yielded 2356 publications and a further 40 were included from reference harvesting and the free search. Duplicate publications were removed (n=173) and following 3 screening phases, 61 publications were eligible and included in the review. Once duplicate reports were collated, a total of 33 original studies were identified and subject to detailed review. Descriptions of these studies can be found in Table 1.

The inter-rater agreement between the reviewers were calculated to be 0.91 for the title screen, 0.64 for the abstract screen and 0.89 the final report; demonstrating very good, good and very good correlation between the reviewers respectively according to Cohen's Kappa coefficient (10).

Figure 1 outlines a PRISMA schematic flow diagram of the review process and reasons for exclusion (11).

Participants and Settings

Across the studies, 1574 participants were included. However, due to inconsistent reporting, overall age and gender information are not available. All participants were considered elderly, and among the studies that reported age information (n= 28; 1411 participants), only 1 participant was under the age of 60. The number of participants included in any given study varied from 3 to 415 subjects. In the 24 studies that reported gender information (comprising 1264 participants), 71% of the participants were women. The majority of studies exclusively assessed participants with a dementia diagnosis (n = 18; 1036 participants), while a further 6 studies (151 participants) included some patients with

dementia. A large proportion of studies were conducted in Japan (n = 10; 178 participants), the US (n = 8; 182 participants), and Australia (n = 4; 577 participants). The most common setting was the nursing home (n = 17; 621 participants). In total, 11 robot systems were used across the studies. Assessed in 22 of the 31 studies, Paro was the most popular choice of SAR intervention. Robots are divided into those capable of learning responses, such as NAO using closed-loop architecture, and those which cannot, such as Paro, using open-loop architecture. In total only 2 closed-loop robots were used (NAO and AIBO) in a total of 6 studies. Descriptions of individual robot systems reviewed can be found in Table 2.

Identified Roles of Socially Assistive Robots

Eligible studies were organised into sets by the role assumed by SAR. Five roles were identified: *Affective Therapy*, *Cognitive Training*, *Social Facilitator*, *Companionship* and *Physiological Therapy*. Specific details of the studies below, such as assessment tools or subject demography, are described in Table 1.

Affective therapy

Fifteen studies (889 participants) evaluated the effect SAR can have in improving the general mood and well-being of elderly participants, or its ability to overcome episodes of mood disturbance. In this review, this role is collectively termed *Affective Therapy*. Nine of these studies (650 participants) were conducted on participants diagnosed with dementia. In total, 11 reported positive findings including reductions in depression scores, agitation scores, and increases in quality of life scores. Whilst these studies were evaluating similar

effects of SAR, their intervention design can broadly be divided into two types: one-on-one interactions with SAR or group interactions with SAR.

Eight studies (657 participants) assessed SAR in one-on-one settings whereas the remaining seven studies (232 participants) had group settings. All of the group setting studies reported positive findings, including reduced agitation and depression levels, and higher expression of positive emotions. Of the 8 one-on-one interaction studies, only 5 report positive findings. Indeed, 2 of these studies(12)(13) report negative findings with increased agitation and worsening dementia, respectively.

These contrasting set of results could indicate a mechanism of how elderly users gain emotional benefit from SAR. A Japanese pilot study(14) assessed group interactions of 26 subjects with Paro and found significant improvements in mood scores during the intervention period. Of note, the authors commented on improved sociability between subjects. As discussed later, several studies(15)(16)(17)(18)(19) demonstrate that SAR can increase the sociability of subjects within groups, which may play a direct role in the mood changes seen here.

Notwithstanding this however, a Dutch crossover study(20) compared two types of one-on-one intervention: therapeutic interventions (Paro introduced at times when subject was distressed) and care support interventions (Paro introduced to facilitate activities of daily living). Only the therapeutic intervention showed a significant improvement in the mood score ($p < 0.01$). This suggests that perhaps while group interventions may be better at generating positive emotions, one-on-one interventions may be appropriate to remedy negative emotions.

Some studies in this set also investigated how SAR compared to soft toys in improving general mood and wellbeing of participants. A large Australian randomised

controlled trial(21) of 415 participants with dementia, compared one-on-one interventions with Paro switched 'on' and Paro switched 'off' (placebo Paro) to identify if Paro's additional social capabilities translated into any positive outcomes. The study found Paro was more effective than usual care in improving pleasure and agitation, but was no different to placebo Paro. Similarly, a Japanese study (8) compared the effect of group interactions with Paro and placebo Paro, and again did not demonstrate any differences between the groups.

These results are mimicked by a Danish randomised controlled trial(13) of 100 subjects, which compared interactions with Paro, a living dog or soft toy cat. The study found intervention type did not affect cognitive state, independence or depression scores and did not affect sleep quality. However, depressive scores improved compared to baseline scores in all groups ($p < 0.05$).

Indeed, only 2 small pilot studies found differences between SAR and soft toys. The first (22) showed subjects engaged more with Paro ($p < 0.05$) and showed more positive emotional expressions with Paro ($p < 0.01$) when compared to a stuffed lion. The second (23) was a study on participants with dementia; it showed that agitation scores were only significantly decreased in a toy cat ($p < 0.05$), whereas NeCoRo (SAR - cat-like robot) only improved scores of pleasure and interest ($p < 0.01$ and $p < 0.05$, respectively).

Cognitive Training

Six studies (344 participants) assessed whether SAR can improve aspects of cognition, such as working memory or executive function, and as such this review has termed this set *Cognitive Training*. This set included 4 studies (239 participants) that assessed elderly subjects with dementia, and 2 studies (105 participants) that assessed

elderly subjects who were cognitively intact. Several robot types have been used in this set including 2 closed loop robots capable of learned responses. This means that whilst broad conclusions surrounding the role of SAR in cognitive training can be made, the evidence for any individual robot system is limited. Five of the six studies (133 participants) concluded with positive findings, although there is a breadth of outcome measures used as surrogate markers for cognitive improvement.

Two studies used cognitive tests, such as Mini-Mental State Examination (MMSE) as the primary outcome measure to assess the impact of SAR interactions. The first was a randomised controlled trial(24) of 34 cognitively healthy subjects in Japan using the Nodding Kabochan as the SAR intervention. Subjects either received the fully functional Nodding Kabochan, or a non-functional Nodding Kabochan (control) for 8 weeks. All interactions were one-on-one with the participant and the SAR in the participants home. Only subjects receiving the functional Nodding Kabochan demonstrated an improved cognitive function score ($p < 0.01$) after the study period. This result contrasts with the conclusion of the previous set, Affective Therapy, where it was difficult to distinguish the positive effects between functional SAR and placebo toys. The distinction here may be that the Nodding Kabochan robot is a communication robot that can talk and sing with the user; a function that a placebo toy is incapable of. The communication itself may be key to this study's findings.

The other study that used cognitive tests as an outcome measure for cognition was a 2-phase block randomised controlled trial(25). This Spanish study involved 101 and 110 subjects with dementia, in the respective phases, and assessed the cognitive effects of group interactions with SAR. In Phase 1, the study compared open-loop system robot, Paro, to closed loop robot, NAO, and a control group, treatment as usual. Compared to control

group, Phase 1 showed a decrease in cognitive function scores in the NAO group only ($p < 0.05$) at follow up. Notably, there were no significant differences between NAO and Paro groups at follow-up. This set of results contrasts with the previous study conducted on cognitively healthy subjects in one-on-one settings. Given different robots systems have been used in the studies, it is difficult to establish which factor is responsible for differing results.

Two studies used neuroimaging modalities as outcome measures of interactions with SAR. The first was a South Korean study(26) that used Magnetic Resonance Imaging (MRI) in a randomised controlled trial of 71 cognitively healthy subjects. The primary outcome measure was change in cortical thickness in brains of participants over the 12 week study period. Subjects were randomised into 3 arms: (1) robot-assisted group training using Silbot and Mero (SAR); (2) traditional intervention training, using computer software; or (3) non-intervention arm - control. The study showed attenuation of cortical thinning on MRI in both intervention groups ($p < 0.05$), and estimated it would take 15.3-months for intervention groups to reach the same level of cortical thinning as controls. This study also used neuropsychiatric tests as a secondary outcome measure. Both intervention groups showed greater improvement in the executive function scores than control group ($p < 0.001$). However, in the general cognitive and visual memory tasks, the traditional intervention group had greater improvement than in the robot group. Indeed, the robot group did not outperform the traditional group on any neuropsychological tests. Both Silbot and Mero are communication robots, like the Nodding Kabochan, which may underpin the improvements in executive function. Nonetheless, the SAR arm did not prove to be any more effective than traditional computer software in either outcome measures for cognitive function.

The other study to use a neuroimaging modality was a Japanese pilot study(27) of 14 subjects with dementia. This study investigated the neuropsychological influence of Paro within an interactive group setting by analysing the electroencephalogram (EEG) recordings. They found an increase in cortical neuronal activity in 7 participants, particularly in participants who liked Paro. It is unclear what the clinical meaning of this finding is, and without a control group, one cannot distinguish the effect of SAR from any other stimulating activity on EEG.

The two final studies used game performance as a surrogate marker for cognitive function in participants with dementia. These were very small studies without control groups. The first(28) included 3 subjects and found that verbal encouragement from SAR (Bandit) improved response time in a game quiz, whilst the second study, with 11 participants, concluded the participants' performance in group ball games and individual card games improved following interactions with SAR (AIBO). Again, the clinical utility of this is unclear, and without objective outcome measures or control groups, there is little that can be learned from these studies.

Social Facilitator

Seven studies (230 participants) assessed the utility of SAR as facilitators for improved sociability between subjects or between subjects and other people. As such, this review has titled this role *Social Facilitator*. All of these studies concluded that the respective SAR intervention improved sociability of participants. Five of these studies (210 participants) were conducted with participants who had been diagnosed with dementia. Four of the studies used Paro as the SAR intervention, and two used AIBO, the robotic dog,

which allowed for a greater degree of comparison between the studies. The final study used Sophie and Jack as the SAR intervention.

Most studies used observed behaviour changes on video recording or via a live assessor during the interaction period. One study(16) used a validated communication scale to assess how group Paro interactions affected sociability. The study concluded that after the 4-week programme, a significant improvement in communication and interaction skills were exhibited by subjects ($p < 0.05$) and an increase in activity participation ($p < 0.05$).

Two studies compared SAR to comparative soft toys/ animals. The first was a crossover study(17) of 23 subjects in the US. Subjects were grouped into sessions with Paro, placebo Paro, or no object. The study concluded that the group with Paro engaged in more social interactions than the group with placebo Paro. This suggests that the sociability effects are associated with SAR itself. The authors note that the novelty around SAR may have contributed to the excitement manifested in increased social engagement. However, as this study was conducted over 4 months, any novelty effects would not likely have been sustained.

The other comparative study was another crossover study(29) in the US, which involved 18 female subjects with dementia. Subjects were divided into sessions with AIBO, a real dog, or no object. The study concluded that although all visit types with AIBO, a dog, or no object, stimulated social interaction by the subject, there were no significant differences in the frequency of social behaviours exhibited by the subjects between visit types.

A similar US pilot study(15) of 7 subjects with dementia was instead conducted in a group setting. Subjects within a group were divided into primary users, those individuals who engaged with Paro at any one time, or non-primary users who were defined as

everyone else in the group. The study showed an increase in social interaction over the 7-week period between primary and non-primary users towards each other and towards staff.

This study's results are reflected in two larger, more recent studies, that also investigate effects of group interactions with SAR on participants with dementia. The first is an Australian study(18) of 139 participants conducted over 5 years with Sophie and Jack. The study reported that social engagement increased over the study period. The second was a Norwegian study(19) with 23 participants, that evaluated the effects of group interactions Paro on those with mild-moderate dementia compared to those with severe dementia. The study found that those with mild-moderate dementia paid more attention to Paro than those with severe dementia. The authors note that SAR interventions may need to be more tailored towards the degree of dementia severity. Another finding was that over the 12-week study period, there was a reported increase in interactions with other subjects, and a decrease in interactions with Paro.

Companionship

Three studies (78 participants) assessed the utility of SAR in overcoming the feeling of loneliness and social isolation in the elderly. These studies are collected into a set this review has titled the *Companionship* role. All 3 of the studies examining SAR in this role showed reductions in loneliness scores. None of these studies were conducted on patients with diagnosed dementia. Two studies used AIBO as the intervention, while the third used Paro.

Only one study assessed this in a one-on-one setting. This was a randomised controlled trial(30) of 38 subjects in the US. Subjects were randomised to have weekly one-on-one sessions with a real dog, AIBO, or no object (control). Subjects in the dog or AIBO group

were significantly less lonely than those in the control group at week 7 ($p < 0.05$ respectively). In both intervention groups, there was a higher attachment score compared to the control group. No significant differences were found between the dog and AIBO groups in the assessment of loneliness, or attachment. This is an important finding that suggests an artificial animal (SAR) can be as effective a companion as a pet.

The other 2 studies were conducted in a group setting. The first study was a pilot study(7) of 11 subjects in Japan using AIBO. Mean loneliness scores after the session were significantly lower than those before the session ($p < 0.05$), although longer term benefits were not established. The second was a larger randomised controlled trial(31) of 34 subjects in New Zealand investigated the effects of Paro on loneliness. Subjects were randomised into a Paro group or a control group that attended normal activities. Subjects in the Paro group had a significantly greater decrease in loneliness score at the 12 week follow-up than the control group ($p < 0.05$). This indicated that sustained effects can be achieved.

The last 2 studies do show promising results, however in the context of the previous set of studies, the decreased sense of loneliness may result from increased sociability in the group setting. Sociability was not measured in either study and therefore may act a confounder.

Physiological Therapy

Two studies (33 participants) investigated the effects of SAR on physiological markers, and as such this review titles this set *Physiological Therapy*. This clinical applicability of this set is less clear, but does raise some questions that future studies may be able to answer. Both of these studies used Paro as the SAR intervention.

The first was a pilot study(32) of 21 subjects in New Zealand and investigated the effect of Paro on blood pressure and heart rate. Subjects had a single 10-minute session with Paro where they were free to interact with the robot. Blood pressure and heart rate was recorded before (T1), immediately after (T2) and 5-minutes after (T3) the 10-minute interaction. Overall, no significant changes in blood pressure or heart rate were demonstrated, however the study decided to exclude 4 residents who did not interact or touch the robot. Subsequently, significant decreases in systolic blood pressure ($p < 0.05$) from T1 to T2 were shown, and such decreases were sustained at T3 measurement. Similarly, significant decreases in diastolic blood pressure ($p < 0.05$) from T1 to T2 were shown, however this decrease was not sustained at T3. Between T1 and T3, heart rate significantly decreased ($p < 0.05$).

In the other study(33) of 12 subjects in Japan, physiological effects of interacting with Paro were investigated. Compared to baseline readings, a significant increase in the ratio of urinary 17-ketosteroid:17-hydroxycorticosteroid ($p < 0.01$), by week 4 of Paro being introduced, was found. The authors suggest this confers an improved physiological reaction to stress. A confounder noted was an increase in social interactions with other residents ($p < 0.05$) by week 4, compared to baseline. It is also not clear from this study if Paro played any role in the increased sociability of residents, however in the context of other studies on the topic, it seems likely.

These two studies do not provide much indication of the clinical use of SAR, however they do give a direction for what future studies could investigate further.

Quantitative Comparison

Several studies reported comparative quantitative data, by using the same or similar assessment scales to others within their role category. The data from these studies have been reproduced from the studies and are compiled in Tables 3-5. As different assessment tools were used across studies, a standardised mean score (0 – 100) was generated to allow comparison across similar assessment tools. Five comparable studies were identified in the Affective Therapy, each using a mood scale to assess either anxiety or depression or both, giving rise to 7 comparable sets of data. Of these, 5 showed significant improvements in the mood scores either in the robot intervention group or in the follow-up score, depending on study design.

Four comparable studies were identified in the Cognitive Training set of studies and of these, 3 showed significant improvements in the cognitive scores. Of note, the 2 phases of the Spanish paper(25) have been listed as two separate sets of data as they are different studies with different interventions and different subject numbers; they both use the same control data however, as seen on Table 4.

Finally, three studies with comparable data were identified in the Companionship set of studies, each of which used validated loneliness scales. All of these studies showed significant improvements in loneliness scores in the robot intervention group or in the follow-up score, depending on study design.

No comparative data was identified in the Social Facilitator or Physiological Therapy groups.

Use of Assistive Robot Technology in Elderly Care

Table 1: Characteristics of Selected Studies

Role	Ref.	Participants	Setting	Intervention/ Study Design	Duration	Measures	Outcome
Affective Therapy	(34)	4 subjects (2 men) aged 82-90, with dementia	Dementia care home, Sweden	Supervised one-on-one interaction with JustoCat. Pilot study.	1 session (Unknown time length)/ week for 7 weeks.	QUALID, CMAI and interview	1. No significant changes observed in scales
	(22)	30 subjects (19 with mild/ moderate dementia + 11 with severe dementia), mean age 84.9 years (mild/ moderate), 87.5 (severe)	Nursing care facility, resident's room, Japan	Supervised one-on-one interaction with Paro and Stuffed Lion. Pilot study.	1 session (~15 mins) for each intervention per subject, separated by 3-6 months	Observed behaviour seen in video-recording	*In both groups:* 1. Subjects talked more frequently to PARO ($p < 0.05$) 2. Showed more positive emotional expressions with PARO ($p < 0.01$) *In Mild/ moderate group only:* 1. Showed more negative emotional expressions with Lion 2. Frequencies of touching and stroking and frequencies of talking to staff member were higher with Lion *In Severe group only:* 1. Showed neutral expression more frequently with Lion
	(20)	71 subjects (14 men) with dementia in 2 groups: therapeutic intervention and care support intervention	Psychogeriatric care institutions, Netherlands	Supervised one-on-one interaction with Paro or no intervention. Paro either served as a therapeutic or a care support tool in two separate phases of the study. Crossover study.	5 sessions (~15 mins)/ month for 2 months; each month of therapy was interspersed with a control month. In the therapeutic arm only, additional sessions were given when patient was in distress.	IPPA and Coop/ Wonca after each interaction	1. Therapeutic-related interventions show an increase of IPPA scores by 2 points ($p < 0.01$) 2. Care support intervention showed no effect
	(35)(36)	53 subjects (20 men) aged 62-95, with a cognitive impairment (MMSE < 25) or diagnosed dementia	Nursing Home, separate room, Norway	Supervised group interaction with Paro or TAU. Randomised controlled trial.	2 sessions (~30 mins)/ week for 12 weeks	Cognitive status, medication, BARS, Norwegian version of CSDD called CDR, QUALID Assessed before (T0), after (T1) and at 3-month follow-up (T2)	1. Reduction in agitation in Paro vs TAU from T0 – T2 ($p < 0.05$) 2. Reduction in depression in Paro vs TAU from T0 – T2 ($p < 0.05$) 3. In those with severe dementia, quality of life scores did not decrease in Paro group from T0 – T2, whereas they did in control. 4. No such difference was found in mild-moderate dementia group
	(37)	18 subjects, aged >65, with dementia	Nursing Home, Australia	Supervised group interaction with Paro or reading group. Randomised controlled trial.	3 sessions (~45 mins)/ week for 5 weeks	Modified QoLAD, RAID, AES, GDS, Revised Algase Wandering Scale–Nursing Home version and OERS	1. The Paro group had higher QOL-AD and OERS-Pleasure scores following the intervention 2. The Paro group had reduced OERS-Anxiety and OERS-Sadness scores following intervention
	(38), (39), (40)(41), (42), (43), (44), (45)	14 subjects (all female) aged 77-	Health Service Facility, Japan	Free group interaction with Paro. Pilot study.	2 sessions (1 hour)/ week for 1 year (and a 5-year	Face scale, GDS and Nursing comments	1. A tendency to improve depression after 8 weeks 2. Improvement in mood

Use of Assistive Robot Technology in Elderly Care

	98, 1 subject without dementia			follow-up)		3. Patients did not lose their interest in the long-term
(13)(46)(47)	100 subjects with a mean age of 85.5	Nursing Homes, Denmark	Supervised one-on-one interaction with Paro, Dog or Toy cat. Randomised controlled trial.	2 sessions (10 mins)/ week for 6 weeks	MMSE, GBS, GDS, CAM, Sleep data and BMI	1. Greater interaction with Paro and Dog compared to toy 2. Cognitive and independence scores worsened over study period in all groups ($p < 0.05$) 3. Depression scores improved over study in all group ($p < 0.05$)
(23)	9 subjects (all female) aged 83-98, with dementia	Nursing Home, US	Supervised one-on-one interaction with NeCoRo and Toy cat. Crossover study.	1 session (10 mins) for each intervention	ABMI, LMBS and observations	1. Both cats maintained participant's interest 2. Significant increase in pleasure ($p < 0.01$) and interest ($p < 0.05$) scores whilst playing with NeCoRo 3. Only the toy cat improved agitation scores ($p < 0.05$)
(14), (48), (49)	26 subjects (all female) aged 73-93, some subjects had dementia	Day service centre, Japan	Free group interaction with Paro. Pilot study.	3 sessions (~45 mins)/ week for 5 weeks	Summarised POMS, Burnout scale for nursing staff, nursing staff comments	1. Significant improvement in POMS scores ($p < 0.05$) 2. Positive social and psychological effects
(8), (50), (51), (52), (53), (54)	23 subjects (6 men) mean age 85	Health Service Facility, Japan	Free group interaction with Paro or placebo Paro. Randomised controlled trial.	4 sessions (1 hour)/ week for 4 weeks	POMS, Face scale, Urinary tests and Nursing comments	1. Improvement in mood and reduction in depression and dejection levels in both groups 2. Urinary results suggest Paro interaction reduces stress
(25)	Phase 1: 20 subjects (10 men), mean age 77.9 Phase 2: 17 subjects (8 men), mean age 79 All subjects were diagnosed with dementia	Day Care Centre, Spain	Phase 1: Supervised group therapy (Cognitive and Physical) with Nao. Phase 2: Supervised group therapy (Cognitive and Physical) with Paro. Crossover study.	2 sessions (30-40mins)/ week for 3 months	GLDS, sMMSE, MMSE, NPI and AI	*Phase 1:* 1. Increase in deterioration scores 2. Significant decrease was seen in irritability scores and total NPI scores *Phase 2:* 1. Increase in deterioration scores
(55)	23 subjects (all men) aged 58 – 97, 19 had been diagnosed with dementia	Veteran Residential Care Facility, US	Supervised one-on-one interaction with Paro. Pilot study.	3 sessions (>5 mins) across 1 year	Behaviour (assessment form designed by authors of study – no formal name) Assessments made before, during and after interaction.	1. Increase in observed positive affective and behavioural indicators (e.g. Bright affect, interacting with others, calm) 2. Decrease in observed negative affective and behavioural indicators (e.g. Anxious, sad, yelling) 3. Those who best responded to Paro were calm and approachable at the before interaction
(21)	415 subjects (101 men) mean age 85. All subjects were diagnosed with dementia	Long term Care Facilities, Australia	Free one-on-one interaction with Paro switched on, Paro switched off or TAU. Cluster-randomised controlled trial.	3 sessions (15 mins)/ week for 10 weeks	Video observations (at baseline and weeks' 1, 5 10 and 15) and CMAI (at baseline and weeks' 10 and 15)	1. Subjects in Paro switched on group were more verbally and visually engaged compared to Paro switched off group 2. Both Paro switched on and switched off groups had reduced neutral affect compared to TAU group 3. Paro switched on was more effective than TAU at improving pleasure and agitation

Use of Assistive Robot Technology in Elderly Care

	(56)	61 subjects (14 men) mean age 84.3 years. All subjects were diagnosed with dementia	Dementia Units, US	Supervised group interaction with Paro or other activity (music, physical activity and mental stimulation). Randomised controlled trial.	3 sessions (20 mins)/ week for 20 weeks	RAID, CSDD, GLDS, pulse rate, pulse oximetry, GSR and medication	1. Anxiety scores, depression scores and pulse rate in Paro group all significantly decreased over the study period compared to control group
	(12)	5 subjects (all female) mean age 84 years. All subjects were diagnosed with dementia	Nursing home, Australia	Supervised one-on-one interaction with CuDDler. Pilot Study.	3 session (30 mins)/ week for 5 weeks	CMAI (before and after each session)	1. Agitation scores increased in 4 of the 5 patients across the 5 week study period
Cognitive Training	(24)	34 subjects (all female), aged > 65 years, living alone	Participant's home, Japan	Living with Nodding Kabochan or control robot (same design as Nodding Kabochan, but cannot talk or nod). Randomised controlled trial.	8 weeks	Questionnaires, BMI, Cognitive tests, APG, and blood and saliva samples	1. Cognitive scores (MMSE + components of Cognistat) were improved in Nodding Kabochan group 2. Saliva cortisol level was decreased in Nodding Kabochan group 3. Higher reports of loss of fatigue, enhancement of motivation and healing in Nodding Kabochan group
	(25)	Phase 1: 101 subjects (13 men), mean age 84.7 Phase 2: 110 subjects (11 men), mean age 84.7 All subjects were diagnosed with dementia	Nursing Home, Spain	Phase 1: Supervised group therapy (Cognitive, Musical and Physical) with Paro or Nao or TAU. Randomised controlled trial. Phase 2: Supervised group therapy (Cognitive, Musical and physical) with Paro or Dog or TAU. Randomised controlled trial.	2 sessions (30-40mins)/ week for 3 months	GLDS, sMMSE, MMSE, NPI, APADEM-NH and the QUALID	*Phase 1:* 1. Decreased apathy in NAO and Paro groups 2. Increased delusions in the NAO group 3. Increased irritability in both robot groups 4. Decrease in scores on the MMSE, but not the sMMSE, in the NAO group 5. There were no significant differences between NAO and Paro groups *Phase 2:* 1. Increase QUALID scores in the Paro group compared to the TAU group 2. Increased hallucinations and irritability in both the Paro and Dog groups compared to the TAU group 3. Increased disinhibition in Paro group compared to Dog group 4. Decreased night-time behaviour disturbances in the Paro group compared to Dog group 1
	(26)	71 healthy subjects, aged >60, based in community	Assessment centre, South Korea	Supervised group interaction with either Silbot and Mero robots (robot cognitive training) or on-screen quiz (traditional cognitive training) or received no cognitive	5 sessions (90 mins)/ week for 12 weeks	MRI, Neuropsychometric tests and Alzheimer's Disease Assessment Scale	1. An attenuation of cortical thinning in both intervention groups 2. Robot therapy showed significantly reduced cortical thinning in the right and left anterior cingulate cortices and small areas of right inferior temporal cortex compared to traditional intervention 3. Global topological organization of white matter

Use of Assistive Robot Technology in Elderly Care

				training (control). Randomised controlled trial.			corticocortical networks was decreased in the control group and the rate of decrease was significantly less in both the intervention groups 4. Robot therapy had greater nodal strength in the left rectus gyrus 5. The intervention groups showed greater improvement in the executive function 6. In the general cognitive and visual memory tasks, the traditional intervention group had greater improvement than in the robot group 7. The robot group did not outperform the traditional group on any neuropsychological test
	(28), (57), (58), (59)	3 subjects (all female) aged >70 with dementia (some reports say 4 subjects, with 1 male)	Care facility, US	Individual interaction (musical, cognitive game) with Bandit (compared to an on-screen simulation of Bandit in some reports). Pilot study.	1 session (20 mins)/ week for 12 months	sMMSE, Response time, Correctness evaluation and Questionnaire	1. Robot encouragement improved response time
	(60)	11 subjects with dementia	Nursing home, Japan	Interaction with AIBO, either individually playing a card game or in a group playing a ball game. Pilot study.	1 session/ day for 5 days	Frequency of activity in video observation	1. Improvement in game performance
	(27), (61)	14 subjects (4 men) mean age 79.2, with dementia	Clinic, Japan	Free group interaction with Paro. Pilot study.	1 session (20 mins)	EEG recording, Questionnaire	1. Improvement in cortical neurons activity of 7 patients, especially in patients who liked the robot
Social Facilitator	(29)	18 subjects (all female) with dementia	Nursing Home, participants room, US	Supervised one-on-one interaction with AIBO, Dog or no object. Crossover study.	1 visit (~ 3 mins)/ week for 3 weeks (each week is a different interaction)	Observed behaviour seen in video-recording	1. All visits generate interactive behaviour with visitor
	(15)	7 subjects with dementia	Dementia rehabilitation wing, US	Supervised group interaction with Paro. Pilot study.	1 session (30-45 mins)/ week for 7 weeks	Observed behaviour of primary and non-primary interactor seen in video-recording	1. PARO increases activity in particular modalities of social interaction, which vary between primary and non-primary interactors 2. PARO improved activity levels
	(16)	12 subjects (9 men), mean age 77.25	Residential care facility, Taiwan	Supervised group interaction with Paro. Pilot study.	2 sessions (30 mins)/ week for 4 weeks	ACIS, Activity Participation scale	1. Significant improvement in communication and interaction skills 2. Significant improvement in activity participation
	(17), (62)	23 subjects, aged 60-104, with high functioning in one nursing home and schizophrenia	Nursing Homes, US	Supervised group interaction with Paro switched on, Paro switched off or no object. Crossover study.	1 session (20 mins)/ 2 weeks (in site A) or per month (in site B) for 4 months (8 sessions vs 4 sessions)	Questionnaire and Observation	1. In switched on Paro group, there was an increase in social interactions; even more in the presence of caregivers or experimenters 2. Switched on Paro also generated feel-good experiences

Use of Assistive Robot Technology in Elderly Care

		and/ or dementia in the other					
	(63)	8 subjects (2 men) aged 68-89, with dementia	Group home, Japan	One-on-one interaction with AIBO. Pilot study.	1 session (30 mins)	N-dementia scale, MMSE, behaviour scale and video observation	1. Improving communication with staff in a group home and establishment of friendly relations with occupants
	(18)	139 subjects (95 men) aged from 65 – 90, with dementia	Residential care facilities, Australia	Supervised group interaction with Sophie and Jack. Observational study.	2 sessions (4-6 hours) across 5 years	Behaviour (assessment form developed by authors –no formal name). Assessments made every 5 minutes during session.	1. Increase in social engagement of subjects across the 5-year study period
	(19)	23 subjects (7 men) aged from 62 – 92. All subjects had a dementia diagnosis	Nursing homes, Norway	Supervised group interaction with Paro. Observational study.	2 sessions (30 mins)/ week for 12 weeks	Observed behaviour as seen in video recording	1. Subjects with mild to moderate dementia paid more attention to Paro than those with severe dementia 2. Over the study period there was an increase in interactions with other subjects, and a decrease in interactions with Paro
Companionship	(30)	38 subjects	Nursing Home, US	Free one-on-one interaction with AIBO/ dog or no object. Randomised controlled trial.	1 session (30 mins)/ week for 8 weeks	Modified LAPS, UCLA LS	1. Dog and AIBO therapy equally reduced loneliness compared to control (more improvement in most lonely participants; in the control group, the most lonely became more lonely) 2. Residents became significantly and equally attached to AIBO and dog. 3. Attachment was not the mechanism for reduced loneliness in dog or AIBO therapy
	(31)(64)	34 subjects, aged >55 years	Retirement Home, New Zealand	Group or individual interaction with Paro or alternative activity. Randomised controlled trial.	2 sessions (1 hour)/ week for 12 weeks	UCLA LS, GDS, QoLAD, interview questionnaire and observations	1. Loneliness scores significantly decrease in the Paro group compared to control 2. Residents enjoyed sharing, interacting and talking about Paro
	(7)	6 subjects (1 man) aged >64 5 separate control subjects used for CgA measurement	Nursing home/ Participant's home, Japan	Free interaction with AIBO. Control group for CgA measurements had no intervention. Pilot study.	4 sessions (1 hour)/ week for 7 weeks	Scores of emotional words, Amount of speech and Satisfaction, AOKLS, SF-36 and salivary CgA	1. Significant reduction of loneliness 2. Improvement in health related quality of life 3. Decrease in salivary CgA, an indicator of sympathetic adrenal system activity 4. Increase in emotional words, amount of speech and satisfaction exhibited
Physiological Therapy	(32)	21 subjects (7 men) mean age 84.9	Residential care facility, New Zealand	Supervised one-on-one interaction with Paro. Pilot study.	1 session (10 mins)	Blood pressure reading: Before during and after interaction	1. Significant reductions in systolic and diastolic blood pressure 2. Reduced systolic blood pressure was sustained after Paro was taken away 3. Reduced diastolic blood pressure was not sustained after Paro was taken away

Use of Assistive Robot Technology in Elderly Care

						4. Data suggests average heart rate decreased
(33), (65), (66), (67), (68), (69)	12 subjects, aged 67-89, with mixed cognitive function	Residential care facility, Japan	Free individual/ group interaction with Paro. Pilot study.	1 session (9.5 hours)/ day for 4 weeks	Urinary tests, interviews and video recording observation	1. Increase in social interaction and density of social networks 2. Improvement of subjects' vital organs reaction to stress

ABMI, Agitated Behaviours Mapping Instrument; ACIS, Assessment of Communication and Interaction Skills; AES, Apathy Evaluation Scale; AI, Apathy Inventory; AIBO, Artificial Intelligence Robot; AOKLS, Ando Osada and Kodama Loneliness Scale; APADEM-NH, Apathy Scale for Institutionalized Patients with Dementia Nursing Home version; APG, Accelerated Plethysmography; BARS, Brief Agitation Rating Scale; BMI, Body Mass Index; CAM, Confusion Assessment Method; CDR, Clinical Dementia Rating Scale; CgA, Chromogranin A; CMAI, Cohen Mansfield Agitation Inventory; Coop/ Wonca, Mood scale; CSDD, Cornell Scale for Symptoms of Depression in Dementia; GBS, Gottfries-Bråne-Steen scale; GDS, Geriatric Depression Scale; GLDS, Global Deterioration Scale; GSR, Galvanic Skin Response; IPPA, Goal attainment scale; LAPS, Lexington Attachment to Pets Scale; LMBS, Lawton's Modified Behaviour Stream; MMSE, Mini Mental State Examination; NAO, NPI, Neuropsychiatric Inventory; OERS, Observed Emotion Rating Scale; POMS, Profile of Mood States; QoLAD, Quality of Life in Alzheimer's Disease Scale; QUALID, Quality of Life Scale; RAID, Rating Anxiety in Dementia Scale; SF-36, Short Form Health Survey; sMMSE, Severe Mini Mental State Examination; TAU, Treatment As Usual; UCLA LS, University of California Los Angeles Loneliness Scale

Use of Assistive Robot Technology in Elderly Patients

Robot	Description	Number Used In Respective Roles					
		Affective Therapy	Cognitive Training	Social Facilitation	Companionship	Physiological Therapy	Total
AIBO	A non-verbal, dog-like robot with a metallic appearance and the ability of sight, walking and interpreting commands. AIBO can learn, mature and, on human interaction, express emotional responses.	-	1	2	2	-	4
Bandit	A humanoid robot mounted on a wheeled base. Bandit can speak, gesticulate and make facial expressions.	-	1	-	-	-	1
CuDDler	A robotic teddy bear able to move its neck, arms and eyelids. CuDDler moves its limbs and vocally interacts. CuDDler can respond appropriately to the pattern and type of touch.	1	-	-	-	-	1
Jack and Sophie	Sophie and Jack are communication robots that are capable facial recognition, emotion recognition, vocalisation, gestures, emotive expressions, singing and dancing.	-	-	1	-	-	1
JustoCat	A non-verbal, cat-like robot with replaceable fur and similar proportions and weight to a real cat. JustoCat is capable of breathing, purring and meowing and is designed to sit on a persons lap and respond to stroking.	1	-	-	-	-	1
Mero	A humanoid head mounted on a base, capable of head motion, facial expressions and speech.	-	1	-	-	-	1
NAO	A humanoid robot, 58 cm tall, capable of walking, speech, gesticulation and dance. NAO is able to interact with people and can develop new skills and become personalised.	1	1	-	-	-	2
NeCoRo	A non-verbal, cat-like robot designed to move and look like a real cat. NecoRo can interpret its surroundings and move accordingly. NeCoRo can express emotion.	1	-	-	-	-	1
Nodding Kabochan	A small robot, with the appearance of a child-like teddy, that can talk, sing and nod. It is designed to communicate with users. Nodding Kabochan can play exercise and singing games with the user.	-	1	-	-	-	1
Silbot	A penguin-like robot that can speak and detect faces. Silbot can engage with users in conversation, games and provide care through drug regimen reminders.	-	1	-	-	-	1
Paro	A non-verbal, seal-like robot with the ability to move its head and tail, blink and make sounds and has 5 sensory modalities: light, sound, temperature, posture and tactile. Paro will respond to being held or stroked and can learn to respond to its name. Paro has its own rhythms; will at times be playful, and at	9	2	3	1	2	17

other times sleepy and inactive.

Table 2: Description of Socially Assistive Robots used in Included Studies

Use of Assistive Robot Technology in Elderly Patients

Table 3: Data extracted from comparable studies in Affective Therapy studies

AFFECTIVE THERAPY Mood Scores									
			Control			**Intervention**			
Study	Number of subjects	Outcome Scale	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	p-value
(34)	4	CMAI	-	-	-	12.6 (6.3)	13.3 (6.6)	0.7	0.88[a]
(35)	53	BARS	22 (19)	23.3 (22)	1.3	20.1 (12.8)	13.7 (11.7)	-6.4	0.044[b]
(35)	53	CSDD	18.2 (12.3)	24.5 (17.3)	6.3	23.7 (12.9)	18.9 (16.8)	-4.8	0.019[b]
(56)	61	CSDD	-	-	-2.1	-	-	-7.4	0.001[b]
(56)	61	RAID	-	-	-0.7	-	-	-3.1	0.003[b]
(37)	18	GDS	-	28.7 (23.3)	-	-	31.3 (19.3)	-	0.72[c]
(13)	100	GDS	-	-	-	13.3 (6.7; 33.3)*	13.3 (6.7; 23.3)*	-	< 0.05[d]

a – Study compares mean baseline score in intervention group to mean follow-up score in the intervention group

b – Study compares change in mean score from baseline to follow-up in control group to change in mean score from baseline to follow-up in intervention group

c – Study compares mean follow up score of control group to mean follow up score of intervention group

d – Study compares median baseline score in intervention group to median follow-up score in the intervention group

* Median and interquartile range reported

BARS, Brief Agitation Rating Scale; CMAI, Cohen Mansfield Agitation Inventory; CSDD, Cornell Scale for Symptoms of Depression in Dementia; GDS, Geriatric Depression Scale; RAID,

Use of Assistive Robot Technology in Elderly Patients

Table 4: Data extracted from comparable studies in Cognitive Training studies

COGNITIVE TRAINING Cognition Scores									
			Control			**Intervention**			
Study	Number of subjects	Outcome Scale	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	p-value
(24)	34	MMSE	-	-	-	94.0 (5)	99.0 (2.3)	5	< 0.01[a]
(25) Phase 1	101	MMSE	12.1 (18.1)	10.4 (15.7)	-1.7	11.8 (17.3)	8.1 (15.0)	-3.7	0.022[b]
(25) Phase 2	110	MMSE	12.1 (18.1)	10.4 (15.7)	-1.7	10.7 (16.5)	9.1 (15.7)	-1.6	0.282[b]
(26)	71	ADAS-Cog	-	-	-	89.9 (5.1)	92.6 (4.0)	2.7	<0.001[a]

Use of Assistive Robot Technology in Elderly Patients

a – Study compares mean baseline score in intervention group to mean follow-up score in the intervention group

b – Study compares change in mean score from baseline to follow-up in control group to change in mean score from baseline to follow-up in intervention group

ADAS-Cog, Alzheimer's Disease Assessment Scale – Cognitive subscale; MMSE, Mini Mental State Examination

Use of Assistive Robot Technology in Elderly Patients

Table 5: Data extracted from comparable studies in Companionship studies

COMPANIONSHIP Loneliness Scores									
			Control			**Intervention**			
Study	Number of subjects	Outcome Scale	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score (SD)	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score (SD)	p-value
(30)	38	UCLA LS	-	-	5.7 (1.3)	-	-	-6.0 (2.7)	< 0.05[b]
(31)	34	UCLA LS	-	-	3.8 (10.3)	-	-	-9.0 (12.6)	0.03[b]
(7)	5	AOKLS	-	-	-	3.3 (2.2)	1.0 (1.3)	-	<0.05[a]

a – Study compares mean baseline score in intervention group to mean follow-up score in the intervention group

b – Study compares change in mean score from baseline to follow-up in control group to change in mean score from baseline to follow-up in intervention group

AOKLS, Ando Osada and Kodama Loneliness Scale; UCLA LS, University of California Los Angeles Loneliness Scale.

DISCUSSION

The aim of this review is to identify the roles SAR could play in elderly care. Despite the infancy of this field, the qualitative amalgamation of the studies demonstrated 5 roles for SAR.

Evaluation of SAR Technology

This review identifies 5 roles for SAR in elderly care: Affective Therapy, Cognitive Training, Social Facilitation, Companionship, and Physiological Therapy. These roles provide a comprehensive classification of how this technology has been utilised in social and physical care to date.

The first set of studies demonstrated that SAR can be used to improve the overall sense of well-being of users, and alleviate acute states of mood disturbance. Interestingly, interactions conducted in a group setting proved to be more consistently effective than one-on-one interactions. However, a study(20) showed that one-on-one interventions were useful in alleviating states of distress. This result may apply to patients with delirium and future studies are required to explore this possibility. The overall picture suggests that whilst SAR is capable of improving mood of subjects, it does not seem to be much better than a comparative soft toy or placebo robot. This is demonstrated in patient groups with and without dementia.

This was not true for the second set, Cognitive Training, where communication robots were significantly more effective at improving cognitive outcome measures than soft toys. The clearest evidence for SAR in improving cognitive function was found in those who are cognitively healthy. Whilst positive findings have been found in participants with dementia, obscure outcome measures make it difficult to interpret the meaning of the

findings. The South Korean study(26) showed that computer programmes are at least as effective as SAR interventions, and may raise doubts about the cost-effectiveness of using SAR to only improve elderly users cognitive function.

All the studies in the Social Facilitator set demonstrated improved sociability. This is demonstrated in subjects with and without dementia, and across 3 robot systems (AIBO, Paro and Sophie and Jack). When compared in group settings, SAR was shown to be more effective than a comparator, such as a soft toy. In one US study (29), subjects were divided into one-on-one sessions with AIBO, a real dog, or no object at all, and whilst all sessions increased frequency of exhibited social behaviour, the study concluded no significant differences between session type. Conversely, in a different US study (17), participants had group interactions with Paro, placebo Paro, or no object. The study concluded that the group with Paro engaged in more social interactions than the group with placebo Paro. This suggests that the sociability effects are associated with a group setting, and perhaps in the absence of a group of users, these effects may not exist.

The Companionship set all showed positive findings. However, 2 studies were conducted in group settings, and the observed improved loneliness scores may be confounded by the increased sociability seen in aforementioned studies. This set has far fewer studies than the other sets generated in this review, however, the findings are insightful. If animal-like SAR can be as much a companion as a pet, then such technology may have particular utility in care homes, where health and safety concerns regarding pets, such as allergies and infection risks, restrict their use.

The final set, Physiological Therapy, did show positive findings, however are clinically uninterpretable. Nonetheless, these studies create new questions about the use of SAR for future studies to address. For example, one study (32) demonstrated short term reductions

in blood pressure and heart rate following Paro interactions. The potential implications of these results are two-fold: this short-term reduction in cardiovascular markers could reflect results seen in the Affective Therapy set, that show calming effects of Paro. Additionally, it may be the case that these reductions can be sustained for the long-term and that SAR may have a role as a non-pharmacological intervention for hypertension. Future studies may benefit from incorporating blood pressure and heart rate outcome measures, alongside other metrics in longer-term studies.

Whilst the utility of SAR in Affective Therapy or Cognitive Training can be replaced by cheaper, existing alternatives (e.g. soft toys or computer software), the main value of SAR may lie in its multi-domain functionality. This review has identified 5 such domains where a single intervention may be of simultaneous value.

Quality of Selected Studies

Of all 33 included studies, 11 were randomised controlled trials (RCTs), 12 included more than 30 subjects and 16 had a comparative intervention. These metrics are not in their own right indicative of the quality of the studies, however together they do provide a general picture. The quality of studies is not evenly distributed across the set. Of the RCTs, 6 are in the Affective Therapy set, while there are none in the Social Facilitator set. Similarly, 9 studies in the Affective Therapy set have a comparative intervention compared to 2 in the Social Facilitator set.

This review did not exclude studies based on methodology. The rationale is that low-quality studies can offer an insight into the potential utility of SAR, and guide study design improvements for future studies. For example, a Companionship role is a popular concept for SAR among commentators in the literature, but very few studies demonstrating this

have been conducted. Evidence supporting a Companionship role is socially desirable because of its applicability to serve the elderly population. As reported by one of the selected studies(30), AIBO, the robotic dog, was as effective a companion as a real dog. This has real implications for its use, specifically where a real animal companion may be inappropriate.

Although no studies were excluded on the basis of quality, there are several underlying methodological limitations facing the selected studies that need to be addressed. Low quality data complicates the task of establishing clinical applications of SAR. It also risks undermining the field's efforts, or sensationalising exploratory research. Another limitation is the narrow set of robots assessed, primarily Paro. This restricts the applicability of results to wider SAR systems with different functionality.

There is also a concern for cultural bias as around a third of the studies were conducted in Japan alone. Although more recent studies have been conducted in other cultural environments, most notably the US and Australia, it is not clear if the results are universally applicable. Additionally, there is evidence of gender bias. Around two-thirds of the participants were women. This is a concern since men and women as populations have been shown to regard robot technology differently(70), and therefore some of the reported findings may be exaggerated, or diminished by the participant composition.

Another common study design issue relates to the supervision of interactions that are present in 20 of the included studies. Although supervision ensures safety for the user, it risks altering how the participant interacts with the robot and may change how the participant reports the robot's utility; known as the Hawthorn Effect. Whilst this is difficult to control for when the study is not randomised and no comparator is used, direct supervision may lead to subjects reporting greater positive effects than is necessarily the

case. An example where this may be the case is a US study (29) where subjects were divided into supervised sessions with AIBO, a real dog, or no object at all. One would anticipate that sessions with an object (AIBO or a soft toy) would stimulate a greater behavioural response than no object at all. However, the study concluded there were no significant differences between the responses to the sessions, irrespective of whether an object was present or not. This suggests that the positive findings were completely independent of the intervention, and may instead be a consequence of supervision.

Another main limitation of the selected studies is the nature of chosen outcome measures. They are often abstract, with a limited number of studies identifying a direct clinical need or problem. Although around half of the studies included a comparator intervention, it often involved uninspiring activities or no activity at all. This is an unfair comparison and may inflate the value attributed to the results. As momentum grows behind SAR, these study design flaws will need to be addressed if the technology is going to play a clinical role in the future.

Review Limitations

The primary limitation of this review is the validity of the categorisation of studies into the defined roles. The roles were created retrospectively, as part of a discovery process on extracting data from the final set of studies. Whilst they have utility in evaluating the state of the field and providing defined expectations for the technology, they have generalised sets of studies that are very different in quality, design and sometimes outcome. There is also the issue that some studies demonstrated several roles for SAR. The studies were categorised on the basis of the the primary outcome measures, irrespective of whether a secondary outcome measure would fit into another set. A consequence of this is that the

weight of data in the respective roles may be misleading. All outcomes have been reported in Table 1 for purposes of data transparency.

Furthermore, this review has an inadvertent risk of excluding relevant papers in the screening phase. Although high concordance between the reviewers was reported, the large volume of studies that had to be reviewed invites the possibility that relevant publications were excluded. The main reason for the high exclusion rate was because the broad search criteria identified irrelevant robot interventions, such as surgical robots or telecommunication devices. It is unlikely, however, that an additional study would have changed the conclusions of this review.

Finally, the comparison of assessment values between studies illustrated in Tables 3-5, aimed to provide some comparison between studies where different outcome measures were used. The comparison does have limitations, because although each assessment tool was scaled from 0-100, a score of 50 in one measure does not necessarily correlate to 50 in a different scale. This has made it difficult to reach broad conclusions about the sets of studies.

Future of the Field

In order to achieve successful application of SAR in elderly care, future studies should be more conscious of the outcome measure chosen and its translation into care. Some studies used surrogate measures such as frequency of laughter(22), or performance in particular games(60). Whilst these may be desired outcomes, it is not clearly demonstrated how they meet quantifiable needs of the elderly population. It is likely that any application of SAR will incorporate several of the previously defined roles. Therefore larger studies should assess the intervention's impact in the context of these clear roles with validated outcome measures. For example, one study(24) involved a robot staying at home with the elderly

participants for 8 weeks, and assessed its impact using questionnaires, cognitive tests, blood and saliva samples. Whilst the study demonstrated an improvement in cognitive scores and a reduction in saliva cortisol, it did not assess whether living with a robot for 8 weeks had any impact on loneliness. Larger randomised controlled trials using valid comparators are needed to definitively show where SAR is and is not useful in elderly care.

Conclusion

Socially assistive robots have shown potential in elderly care which, in light of recent demographic shifts, promises to reform the delivery of care for the elderly. Although many of the studies described have methodological issues, the size and quality of studies are improving. This review has qualitatively assessed the existing research and comprehensively outlined the state of the field as it stands. In establishing the 5 roles to which SAR can be ascribed, this review intends not to restrict ambition, but to provide a basis for clinical applicability and design of future studies. This review urges that new studies should be clearer about the precise role any robot intervention intends to serve, and use validated measures to assess their effectiveness. Future studies need to demonstrate how SAR can solve real problems in order to shift from novelty to functionality in elderly care.

Acknowledgements

This review would like to thank Peter Forrest for his scientific guidance and participation in discussions for this project, and Sundhiya Mandalia for her support in interpreting of the data and its analysis.

Declaration of competing interests: None

Contribution (Author contributing):

Design of the review (Jordan Abdi, Marcela Vizcaychipi)
Data collection (Jordan Abdi, Ahmed Al-Hindawi)
Data analysis and interpretation (Jordan Abdi, Tiffany Ng, Marcela Vizcaychipi)
Drafting the article (Jordan Abdi, Ahmed Al-Hindawi, Tiffany Ng)
Critical revision of the article (Jordan Abdi, Tiffany Ng, Marcela Vizcaychipi)
Final approval of the version to be published (Marcela Vizcaychipi)

Funding sources: All authors declare independence from funders.

Transparency Declaration: All authors, external and internal, had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. The lead author affirms the manuscript is honest, accurate and that any discrepancies from the study are accounted for.

Data sharing: no additional data available.

Systematic registration number and name of register: NIHR 58672

References

1. United Nations. World Population Prospects. 2015;1. Available from: http://esa.un.org/unpd/wpp/Publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf
2. Thomas J. Insights into Loneliness , Older People and Well-being. 2015;1–10. Available from: http://backup.ons.gov.uk/wp-content/uploads/sites/3/2015/10/Insights-into-Loneliness-Older-People-and-Well-being-2015.pdf
3. Victor C, Scambler S, Bond J, Bowling A. Loneliness, social isolation and living alone in later life. Rev Clin Gerontol. 2000;10(4):407–17.
4. Holt-lunstad J, Smith TB, Layton JB. Social Relationships and Mortality Risk : A Meta-analytic Review. PLoS Med. 2010;7(7).
5. Hegel F, Muhl C, Wrede B, Hielscher-Fastabend M, Sagerer G. Understanding Social Robots. Second Int Conf Adv Comput Interact. 2009;(Section II):169--174.
6. Reiser U, Jacobs T, Arbeiter G, Parlitz C. Care-O-bot ® 3 – Vision of a Robot Butler. 2013. 97-116 p.
7. Kanamori M, Suzuki M, Oshiro H. Pilot study on improvement of quality of life among elderly using a pet-type robot. IEEE Int Conf Robot Autom. 2003;107–12.
8. Wada K, Shibata T, Saito T, Tanie K. Effects of Robot Assisted Activity to Elderly People who Stay at a Health Service Facility for the Aged. IEEE/RSJ Int Conf Intell Robot Syst. 2003;(October).
9. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. 2011. Available from: http://handbook.cochrane.org
10. Orwin RG, Vevea JL. Evaluating Coding Decisions. In: The Handbook of Research Synthesis and Meta-Analysis. 1994. p. 177–200.

Use of Assistive Robot Technology in Elderly Patients

11. Moher D, Liberati A, Tetzlaff J, Altman DG, Grp P. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement (Reprinted from Annals of Internal Medicine). PLoS Med [Internet]. 2009;89(9):873–80. Available from: <Go to ISI>://000270260200002

12. Moyle W, Jones C, Sung B, Bramble M. What Effect Does an Animal Robot Called CuDDler Have on the Engagement and Emotional Response of Older People with Dementia ? A Pilot Feasibility Study. Int J Soc Robot [Internet]. Springer Netherlands; 2016;8(1):145–56. Available from: "http://dx.doi.org/10.1007/s12369-015-0326-7

13. Thodberg K, Sorensen LU, Christensen JW, Poulsen PH, Houbak B, Damgaard V, et al. Therapeutic effects of dog visits in nursing homes for the elderly. [Internet]. Psychogeriatrics. United Kingdom: Wiley-Blackwell Publishing Ltd.; 2015. p. 289–97. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc11&NEWS=N&AN=2015-49600-001

14. Wada K, Shibata T, Saito T, Tanie K. Effects of robot assisted activity for elderly people at day service center and analysis of its factors. Proc 4th World Congr Intell Control Autom (Cat No02EX527) [Internet]. 2002;2:1301–5. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1020792

15. Sabanovic S, Bennett CC, Chang W-L, Huber L. PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. IEEE Int Conf Rehabil Robot [Internet]. United States; 2013;2013:6650427. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24187245

16. Sung H-C, Chang S-M, Chin M-Y, Lee W-L. Robot-assisted therapy for improving social interactions and activity participation among institutionalized older adults: a pilot study. Asia

Pac Psychiatry [Internet]. Australia; 2015;7(1):1–6. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24692085

17. Kidd CD, Taggart W, Turkle S. A Sociable Robot to Encourage Social Interaction among the Elderly. IEEE Int Conf Robot Autom. 2006;3972–6.

18. Chu M, Khosla R, Mohammad S, Khaksar S. Service innovation through social robot engagement to improve dementia care quality. 2017;435(September).

19. Jøranson N, Pedersen I, Marie A, Rokstad M, Aamodt G, Olsen C, et al. Group activity with Paro in nursing homes : systematic investigation of behaviors in participants. 2017;(2016):1345–54.

20. Bemelmans R, Gelderblom GJ, Jonker P, de Witte L. Effectiveness of Robot Paro in Intramural Psychogeriatric Care: A Multicenter Quasi-Experimental Study. J Am Med Dir Assoc [Internet]. United States; 2015;16(11):946–50. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=26115817

21. Moyle W, Jones CJ, Mur JE, Thalib L, Beattie ERA, Shum DKH, et al. Use of a Robotic Seal as a Therapeutic Tool to Improve Dementia Symptoms : A Cluster-Randomized Controlled Trial. 2017;18:766–73.

22. Takayanagi K, Kirita T, Shibata T. Comparison of Verbal and Emotional Responses of Elderly People with Mild/Moderate Dementia and Those with Severe Dementia in Responses to Seal Robot, PARO. Front Aging Neurosci [Internet]. Switzerland; 2014;6:257. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=25309434

23. Libin A, Cohen-Mansfield J. Therapeutic robocat for nursing home residents with dementia: preliminary inquiry. Am J Alzheimers Dis Other Demen [Internet]. United States;

2004;19(2):111–6. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med5&NEWS=N&AN=15106392

24. Tanaka M, Ishii A, Yamano E, Ogikubo H, Okazaki M, Kamimura K, et al. Effect of a human-type communication robot on cognitive function in elderly women living alone. Med Sci Monit [Internet]. Poland; 2012;18(9):CR550-7. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=22936190

25. Valentí Soler M, Agüera-Ortiz L, Olazarán Rodríguez J, Mendoza Rebolledo C, Pérez Muñoz A, Rodríguez Pérez I, et al. Social robots in advanced dementia. Front Aging Neurosci [Internet]. Frontiers; 2015 Jan 3 [cited 2016 Mar 14];7:133. Available from: http://journal.frontiersin.org/article/10.3389/fnagi.2015.00133/abstract

26. Kim GH, Jeon S, Im K, Seo SW, Cho H, Noh Y, et al. Structural brain changes after robot-assisted cognitive training in the elderly: A single-blind randomized controlled trial. Alzheimer's Dement [Internet]. 2013 Jul [cited 2016 Mar 14];9(4):P476–7. Available from: http://www.sciencedirect.com/science/article/pii/S1552526013016282

27. Wada K, Shibata T, Musha T, Kimura S. Robot Therapy for Elders Affected by Dementia. 2014 IEEE Long Isl Syst Appl Technol Conf LISAT 2014. 2014;(August).

28. Tapus A. Improving the quality of life of people with dementia through the use of socially assistive robots. Proc - 2009 Adv Technol Enhanc Qual Life, AT-EQUAL 2009. 2009;81–6.

29. Kramer SC, Friedmann E, Bernstein PL. Comparison of the effect of human interaction, animal-assisted therapy, and AIBO-assisted therapy on long-term care residents with dementia. [Internet]. Anthrozoos. Berg Publishers; 2009. p. 43–57. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc6&NEWS=N&AN=2009-

04201-004

30. Banks MR, Willoughby LM, Banks WA. Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. J Am Med Dir Assoc [Internet]. United States; 2008;9(3):173–7. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med5&NEWS=N&AN=18294600

31. Robinson H, Macdonald B, Kerse N, Broadbent E. The psychosocial effects of a companion robot: a randomized controlled trial. J Am Med Dir Assoc [Internet]. United States; 2013;14(9):661–7. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=23545466

32. Robinson H, MacDonald B, Broadbent E. Physiological effects of a companion robot on blood pressure of older people in residential care facility: a pilot study. Australas J Ageing [Internet]. Australia; 2015;34(1):27–32. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24373064

33. Wada K, Shibata T. Social and physiological influences of robot therapy in a care house. [Internet]. Special Issue: Human and robot interactive communication. 2008. p. 258–76. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc6&NEWS=N&AN=2008-08717-006

34. Gustafsson C, Svanberg C, Müllersdorf M. Using a Robotic Cat in Dementia Care. J Gerontol Nurs [Internet]. Senior Lecturer, School of Health, Care, and Social Welfare, Mälardalen University, Eskilstuna and Västerås, Sweden: SLACK Incorporated; 2015 Oct;41(10):46–56 11p. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=110074731&site=ehost-

live

35. Joranson N, Pedersen I, Rokstad AMM, Ihlebaek C. Effects on Symptoms of Agitation and Depression in Persons With Dementia Participating in Robot-Assisted Activity: A Cluster-Randomized Controlled Trial. J Am Med Dir Assoc [Internet]. United States; 2015;16(10):867–73. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=26096582

36. Jøranson N, Pedersen I, Marie A, Rokstad M, Ihlebæk C. Change in quality of life in older people with dementia participating in Paro-activity : a cluster-randomized controlled trial. 2016;(June).

37. Moyle W, Cooke M, Beattie E, Jones C, Klein B, Cook G, et al. Exploring the effect of companion robots on emotional expression in older adults with dementia: a pilot randomized controlled trial. J Gerontol Nurs [Internet]. United States; 2013;39(5):46–53. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=23506125

38. Wada K, Shibata T, Saito T, Sakamoto K, Tanie K. Psychological and Social Effects of One Year Robot Assisted Activity on Elderly People at a Health Service Facility for the Aged. Proc IEEE Int Conf Robot Autom. 2003;(April):2785–90.

39. Wada KWK, Shibata TST, Saito TST, Sakamoto KSK, Tanie KTK. Robot assisted activity at a health service facility for the aged for 17 months: an interim report of long-term experiment. IEEE Work Adv Robot its Soc Impacts, 2005. 2005;127–32.

40. Wada K, Shibata T, Sakamoto K, Tanie K. Quantitative analysis of utterance of elderly people in long-term robot assisted activity. Proc - IEEE Int Work Robot Hum Interact Commun. 2005;2005:267–72.

41. Wada K, Shibata T, Saito T, Tanie K. Psychological and social effects in long-term experiment

of robot assisted activity to elderly people at a health service facility for the aged. 2004 IEEE/RSJ Int Conf Intell Robot Syst [Internet]. 2004;3:3068–73. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-14044273830&partnerID=40&md5=b85e82f0c643a517d76b6eb7a05fe686

42. Wada K, Shibata T, Kawaguchi Y. Long-term robot therapy in a health service facility for the aged-A case study for 5 years. 2009 IEEE Int Conf Rehabil Robot ICORR 2009 [Internet]. 2009;930–3. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-70449380117&partnerID=40&md5=4bba7ff889ec97a3bd7ebd52ab8af615\nhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5209495\nhttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5209495

43. Wada K, Shibata T, Sakamoto K, Saito T, Tanie K. Analysis of utterance in long-term robot assisted activity for elderly people. Proceedings, 2005 IEEE/ASME Int Conf Adv Intell Mechantronics [Internet]. 2005;1(April):31–6. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1500961

44. Wada K, Shibata T, Saito T, Tanie K. Robot assisted activity at a health service facility for the aged for ten weeks: an interim report of a long-term experiment. Int IEEE/EMBS Conf Neural Eng NER. 2006;470–3.

45. Wada K, Shibata T, Saito T, Tanie K. Effects of Three Months Robot Assisted Activity to Depression of Elderly People who stay at a Health Service Facility for the Aged. SICE 2004 Annu Conf. 2004;2709–14.

46. Thodberg K, Sørensen LU, Christensen JW, Poulsen PH, Houbak B, Damgaard V, et al. Therapeutic effects of dog visits in nursing homes for the elderly. 2016;289–97.

47. Thodberg K, Sørensen LU, Videbech PB, Poulsen PH, Houbak B, Damgaard V, et al. Behavioral

Responses of Nursing Home Residents to Visits From a Person with a Dog , a Robot Seal or aToy Cat Behavioral Responses of Nursing Home Residents to Visits From a Person with a Dog , a Robot Seal or aToy Cat. 2016;7936(September 2017).

48. Wada K, Shibata T, Saito T, Tanie K. Analysis of factors that bring mental effects to elderly people in robot assisted activity. IEEE/RSJ Int Conf Intell Robot Syst. 2002;2(October):1152–7.

49. Wada K, Shibata T, Saito T, Tanie K. Robot Assisted Activity for Elderly People and Nurses at a Day Service Center. IEEE/RSJ. 2002;(May):1416–21.

50. Wada K, Shibata T, Saito T, Tanie K. Robot assisted activity to elderly at a health service facility for the aged. Int IEEE/EMBS Conf Neural Eng NER. 2003;

51. Saito T, Shibata T, Wada K, Tanie K. Relationship between interaction with the mental commit robot and change of stress reaction of the elderly. Proc IEEE Int Symp Comput Intell Robot Autom CIRA. 2003;1:119–24.

52. Wada K, Shibata T, Saito T, Tanie K. Relationship between familiarity with mental commit robot and psychological effects to elderly people by robot assisted activity. Proc IEEE Int Symp Comput Intell Robot Autom CIRA. 2003;1:113–8.

53. Wada K, Shibata T, Saito T, Tanie K. Psychological and Social Effects of Robot Assisted Activity to Elderly People who stay at a Health Service Facility for the Aged. IEEE Int Conf Robot Autom. 2004;3:14–9.

54. Wada K, Shibata T, Saito T, Tanie K. Effects of robot-assisted activity for elderly people and nurses at a day service center. Proc IEEE. 2004;92(11):1780–8.

55. Lane GW, Noronha D, Rivera A, et al. Effectiveness of a Social Robot , " Paro ," in a VA Long-Term Care Setting. 2016;(August).

56. Petersen S, Houston S, Qin H, Tague C, Studley J, Health W, et al. The Utilization of Robotic Pets in Dementia Care. 2017;55:569–74.

57. Tapus A, Ţăpuş C, Matarić MJ. The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. 2009 IEEE Int Conf Rehabil Robot ICORR 2009. 2009;924–9.

58. Tapus A, Tapus C, Mataric M. Music therapist robot for individuals with cognitive impairments. ACM/IEEE Int Conf Human-Robot Interact. 2009;297–8.

59. Tapus A, Tapus C, Mataric M. Music Therapist Robot: A Solution for Helping People with Cognitive Impairments. 2009;297–8.

60. Hamada T, Okubo H, Inoue K, Maruyama J, Onari H, Kagawa Y, et al. Robot therapy as for recreation for elderly people with dementia - Game recreation using a pet-type robot -. Proc 17th IEEE Int Symp Robot Hum Interact Commun RO-MAN. 2008;174–9.

61. Wada K, Shibatal T, Musha T, Kimura S. Effects of robot therapy for demented patients evaluated by EEG. 2005 IEEE/RSJ Int Conf Intell Robot Syst IROS. 2005;2205–10.

62. Taggart W, Turkle S, Kidd CD. An Interactive Robot in a Nursing Home : Preliminary Remarks Field Setting : Nursing Homes. Towar Soc Mech Android Sci A COGSCI Work. 2005;1–6.

63. Sakairi K. Research of robot-assisted activity for the elderly with senile dementia in a group home. SICE 2004 Annu Conf. 2004;3:2092–4.

64. Robinson H, Broadbent E, MacDonald B. Group sessions with Paro in a nursing home: Structure, observations and interviews. Australas J Ageing. 2015 Jun;

65. Wada K, Shibata T. Social effects of robot therapy in a care house - Change of social network of the residents for two months. Proc - IEEE Int Conf Robot Autom. 2007;(April):1250–5.

66. Wada K. Social and physiological influences of living with seal robots in an elderly care house for two months. Gerontechnology [Internet]. 2008;235–235. Available from: http://www.gerontechnology.info/Journal/pdf.php?a=645

67. Wada K, Shibata T. Robot therapy in a care house - Change of relationship among the residents

and seal robot during a 2-month long study. Proc - IEEE Int Work Robot Hum Interact Commun. 2007;107–12.

68. Wada K, Shibata T. Robot therapy in a care house - Results of case studies. Proc - IEEE Int Work Robot Hum Interact Commun. 2006;(June 2005):581–6.

69. Wada K, Shibata T. Living With Seal Robots—Its Sociopsychological and Physiological Influences on the Elderly at a Care House. Proc - IEEE Int Work Robot Hum Interact Commun. 2007;23(5):972–80.

70. Schermerhorn P, Scheutz M, Crowell CR. Robot social presence and gender: Do females view robots differently than males? 3rd ACM/IEEE Int Conf. 2008;

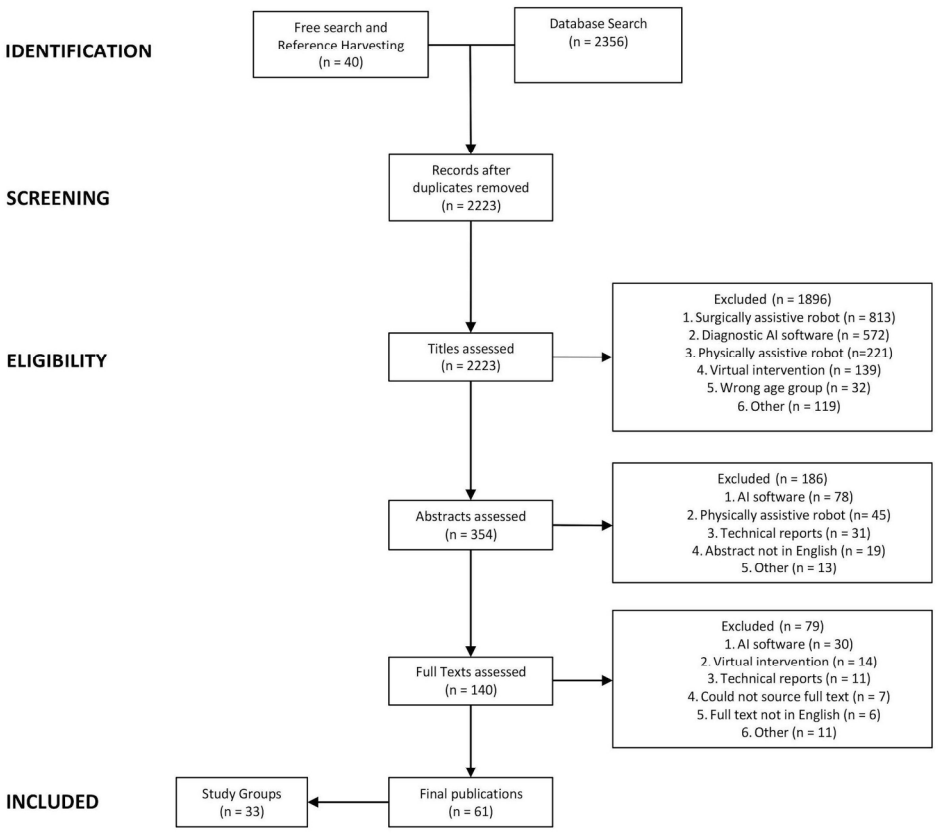

Figure 1: Schematic flow diagram of the review process

179x166mm (300 x 300 DPI)

An example of bibliographical search for PubMEd:

1. Service robot* [Text Word]
2. Therapeutic robot* [Text Word]
3. Socially assistive robot* [Text Word]
4. AIBO [Text Word]
5. Paro [Text Word]
6. Care-o-bot [Text Word]
7. Robotics [MeSH]
8. Artificial Intelligence [MeSH]
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10. Aged, 80 and over [MeSH]
11. Dementia [MeSH]
12. Elder* [Text Word]
13. Senior* [Text Word]
14. Older person [Text Word]
15. Geriatric* [Text Word]
16. Old people [Text Word]
17. 10 or 11 or 12 or 13 or 14 or 15 or 16
18. 9 and 17

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	11
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	11
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	11
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	19
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	10
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	19
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	29
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	31
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	34
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	36

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

BMJ Open

A Scoping Review on the Use of Socially Assistive Robot Technology in Elderly Care

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-018815.R2
Article Type:	Research
Date Submitted by the Author:	10-Dec-2017
Complete List of Authors:	Abdi, Jordan; Imperial College London, Al-Hindawi, Ahmed; Chelsea and Westminster Hospital, Magill Department of Anaesthesia Ng, Tiffany; Chelsea and Westminster Hospital, Magill Department of Anaesthesia Vizcaychipi, M; Imperial College London, Magill Deaprtment Anaesthesia
Primary Subject Heading:	Geriatric medicine
Secondary Subject Heading:	Sociology, Public health, Mental health
Keywords:	Socially Assistive Robots, Elderly, Robots, GERIATRIC MEDICINE, SOCIAL MEDICINE

A Scoping Review on the Use of Socially Assistive Robot Technology in Elderly Care

Jordan Abdi[1], Ahmed Al-Hindawi[2], Tiffany Ng[2], Marcela P Vizcaychipi[2]

[1]Department of Medicine, Imperial College London, Kensington, London SW7 2AZ

[2]Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH

Author note:

Jordan Abdi (Abdi J), Undergraduate student, Department of Medicine, Imperial College London

Ahmed Al-Hindawi (Al-Hindawi A), Intensive Care Registrar, Magill Research Fellow, Magill Department of Anaesthesia, Chelsea and Westminster Hospital

Tiffany Ng (Ng T), Doctor, Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH.

Marcela Vizcaychipi (Vizcaychipi MP), Consultant in Anaesthesia and Intensive Care in the Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH

Address Correspondence: Jordan Abdi, Magill Department of Anaesthesia, Intensive Care Medicine and Pain Management, Chelsea and Westminster Hospital, London SW10 9NH.

Electronic address: ja1612@ic.ac.uk

Use of Assistive Robot Technology in Elderly Care

STRUCTURED ABSTRACT

OBJECTIVE: With an elderly population that is set to more than double by 2050 worldwide, there will be an increased demand for elderly care. This poses several impediments in the delivery of high quality health and social care. Socially Assistive Robot (SAR) technology could assume new roles in health and social care to meet this higher demand. This review qualitatively examines the literature on the use of SAR in elderly care and aims to establish the roles this technology may play in the future.

DESIGN: Scoping Review

DATA SOURCES: Search of CINAHL, Cochrane Library, EMBASE, MEDLINE, PsychINFO and SCOPUS databases was conducted, complemented with a free search using Google Scholar and reference harvesting. All publications went through a selection process, which involved sequentially reviewing the title, abstract and full text of the publication. No limitations regarding date of publication were imposed and only English publications were taken into account. The main search was conducted in March 2016, and the latest search was conducted in September 2017.

ELIGIBILITY CRITERIA: The inclusion criteria consist of elderly participants, any elderly healthcare facility, humanoid and pet robots, and all social interaction types with the robot. Exclusions were acceptability studies, technical reports of robots and publications surrounding physically or surgically assistive robots.

RESULTS: In total, 61 final publications were included in the review, describing 33 studies and including 1574 participants and 11 robots. 28 of the 33 papers report positive findings. Five roles of SAR were identified: Affective Therapy, Cognitive Training, Social Facilitator, Companionship and Physiological Therapy.

CONCLUSIONS: Although many positive outcomes were reported, a large proportion of the studies have methodological issues, which limit the utility of the results. Nonetheless, the reported value of SAR in elderly care does warrant further investigation. Future studies should endeavour to validate the roles demonstrated in this review.

SYSTEMATIC REVIEW REGISTRATION - NIHR 58672

Abstract Word Count: 334 Words

Keywords: Socially assistive robots, elderly, robots, geriatric care, social care

STRENGTHS AND LIMITATIONS

- This is the first scoping review of the literature that has evaluated and categorised the effects of Socially Assistive Robot interventions aimed to improve the health and social care of elderly people.
- The novelty of the field means that the quantity and quality of studies available in the current literature is limited, making generalisations difficult.
- The retrospective creation of SAR roles grouped together sets of studies that differed in quality, design and sometimes outcome, which may mislead the actual weight of data in the respective roles.

INTRODUCTION

The global population is undergoing a demographic shift. Life expectancy is growing and the post-war baby boom generation is entering retirement. The implications on resource allocation will impact the delivery of elderly care. As of 2015(1), 21% of Western Europe's population were over the age of 60, and this is expected to rise to 33% by 2030. By 2050, there are expected to be more people over the age of 60 globally than under 15, reaching a total population of 2.1 billion compared to 901 million in 2015. This is compounded by a proportional decrease in the number of social and health care providers shouldering this increased burden. In 2015, 7 workers were allocated for every elderly person globally, but this is projected to fall to 4.9 in 15-years(1). Moreover, the situation is magnified in Europe by an accelerated ageing population. Currently, there are 3.5 workers for every elderly person, but this is set to fall to 2.4 by 2030. The shift in societal proportions will place new pressures on all aspects of elderly care.

Loneliness, for instance, is a consequence of social, psychological and personal factors. Over half of people over the age of 75 live alone(2) and 17% of older people see family, friends or neighbours less than once a week(3). A recent meta-analysis(4) showed that the impact of loneliness and isolation carries the same mortality risk as smoking 15 cigarettes a day. This is compounded by the fact that social care is a labour intensive industry in a world with a proportionally shrinking workforce.

Throughout many industries the 'robot revolution' promises to solve this growing personnel shortage. At present, physically or surgically assistive robots dominate the healthcare sector's robot usage. This includes: (i) increasingly sophisticated wheelchairs transforming the limitations imposed on paraplegics; (ii) robotic limbs redefining amputee capabilities; and (iii) robotic surgeons revolutionising how and where surgery can be

performed. Nonetheless, physically assistive robots do not combat the increasing mental health burden recognised in the elderly population. It is here that the concept of socially assistive robots is gaining headway. These are robots adept at completing a complex series of physical tasks with the addition of a social interface capable of convincing a user that the robot is a social interaction partner(5).

Socially assistive robots have been categorised into 2 operational groups: 1) service robots and 2) companion robots. Service robots are tasked with aiding activities of daily living(6). Companion robots, by contrast, are more generally associated with improving the psychological status and overall well-being of its users. Such examples include Sony's AIBO(7) and Paro(8). Despite much of the hype, the utilisation of this technology in elderly care is not completely ascertained.

The aim of this scoping review is to establish the clinical usefulness of Socially Assistive Robots (SAR) in elderly care. Through examination and qualitative analyses of existing literature, studies will showcase the utility of SAR and their associated clinical outcomes. A better understanding of SAR and its ability to provide integral care, both socially and physiologically, will provide an indication of its future role in society.

Use of Assistive Robot Technology in Elderly Care

METHODOLOGY

The protocol for this review was conducted in accordance with the principles of the Cochrane Handbook for Systematic Reviews of Interventions (9).

Search Strategy

The following bibliographical databases were searched: CINAHL, Cochrane Library, EMBASE, MEDLINE, PsychINFO and SCOPUS using medical subject headings (MeSH or where appropriate, the database specific thesaurus equivalent) or text word terms. The database search query was composed of 2 search concepts: the intervention (SAR) and the context (elderly care). Free text terms for the intervention included: "service robot*", "therapeutic robot*" and "socially assistive robot*"; their associated MeSH terms were "Robotics" and "Artificial Intelligence". The names of specific robot systems were also searched for. The free words used for the context included: "elder*", "senior*", "older person*", "old people" and "dementia"; their associated MeSH term was "Aged, 80 and over". The use of the asterisk (*) enables the word to be treated as a prefix. For example, "elder*" will represent "elderly" and "eldercare" amongst others. (see supplementary material for an example of a bibliographical search). Additional studies were selected through a free search (Google Scholar) and from reference lists of selected publications and relevant reviews. The main search was conducted in March 2016, and the latest search was conducted in September 2017.

Study Selection

Two reviewers (JA and AA) independently screened the publications in a three-step assessment process: the title, abstract and full text and selection were made in accordance

with inclusion criteria. All publications collected during the database search, free search and reference list harvesting were scored on a 3-point scale (0 = Not relevant, 1 = Possibly relevant, 2 = Very relevant) and those with a combined score of 2 between the reviews would make it through to the next round of scoring. All publications with a total score of 0 were excluded. A publication with a combined score of 1 indicated a disagreement between the reviewers, and would be resolved through discussion. At the end of the full text screening round, a final set of publications to be included into the review was acquired. Cohen's kappa coefficient was calculated to ascertain the agreement between the reviewers in the title, abstract and full text screening phases.

A study was considered eligible if it assessed the usefulness of SAR in the elderly population with a clinical outcome measure. A study that simply assessed the robot's acceptability to elderly users without a clinical outcome measure, or was a technical report, or concerned the use of physically or surgically assistive robots was excluded. No limitations regarding date of publication were imposed and only English publications were considered.

Since the field of Socially Assistive Robotics is in its infancy, many of the studies are small and exploratory. Nonetheless, they provide an insight into what is currently being researched and the potential applications of SAR in elderly care. For this reason, no publication was excluded on the grounds of methodological quality.

Data Extraction

The data extraction form was designed in line with the PICO approach (Participants, Intervention, Comparator and Outcomes). This process was conducted by one reviewer (JA) to ensure consistent extraction of all studies. All clinical outcome measures reported in selected studies were extracted. Data extraction included, in addition to outcomes, country

in which study was conducted, number of included participants, mean age of participants, gender ratio of participants, specific robot used, cognitive status of participants, settings, study design, study duration, and assessment tools.

Duplicate reports of the same study may present in different journals, papers or conference proceedings, and may each focus on different outcome measures or include a follow up data point. To minimise the impact of duplicates, the final set of publications were collated into "study groups" containing duplicate reports. The data extraction process was conducted on the most comprehensive report of a given study.

Data Synthesis and Analysis

Studies were categorised into groups by the role of the robot in the study. The categories were generated retrospectively by the authors and were not pre-defined or directly referenced in the original studies themselves.

Some studies used comparable quantitative outcome measures in their assessment of clinical utility of SAR robots. As different assessment tools were used across studies, a standardised mean score (0 – 100) was generated to allow comparison across similar assessment tools. The result is a unit-free size.

Use of Assistive Robot Technology in Elderly Care

RESULTS

Search Results

The database search yielded 2356 publications and a further 40 were included from reference harvesting and the free search. Duplicate publications were removed (n=173) and following 3 screening phases, 61 publications were eligible and included in the review. Once duplicate reports were collated, a total of 33 original studies were identified and subject to detailed review. Descriptions of these studies can be found in Table 1.

The inter-rater agreement between the reviewers were calculated to be 0.91 for the title screen, 0.64 for the abstract screen and 0.89 the final report; demonstrating very good, good and very good correlation between the reviewers respectively according to Cohen's Kappa coefficient (10).

Figure 1 outlines a PRISMA schematic flow diagram of the review process and reasons for exclusion (11).

Participants and Settings

Across the studies, 1574 participants were included. However, due to inconsistent reporting, overall age and gender information are not available. All participants were considered elderly, and among the studies that reported age information (n= 28; 1411 participants), only 1 participant was under the age of 60. The number of participants included in any given study varied from 3 to 415 subjects. In the 24 studies that reported gender information (comprising 1264 participants), 71% of the participants were women. The majority of studies exclusively assessed participants with a dementia diagnosis (n = 18; 1036 participants), while a further 6 studies (151 participants) included some patients with

dementia. A large proportion of studies were conducted in Japan (n = 10; 178 participants), the US (n = 8; 182 participants), and Australia (n = 4; 577 participants). The most common setting was the nursing home (n = 17; 621 participants). In total, 11 robot systems were used across the studies. Assessed in 22 of the 31 studies, Paro was the most popular choice of SAR intervention. Robots are divided into those capable of learning responses, such as NAO using closed-loop architecture, and those which cannot, such as Paro, using open-loop architecture. In total only 2 closed-loop robots were used (NAO and AIBO) in a total of 6 studies. Descriptions of individual robot systems reviewed can be found in Table 2.

Identified Roles of Socially Assistive Robots

Eligible studies were organised into sets by the role assumed by SAR. Five roles were identified: *Affective Therapy*, *Cognitive Training*, *Social Facilitator*, *Companionship* and *Physiological Therapy*. Specific details of the studies below, such as assessment tools or subject demography, are described in Table 1.

Affective therapy

Fifteen studies (889 participants) evaluated the effect SAR can have in improving the general mood and well-being of elderly participants, or its ability to overcome episodes of mood disturbance. In this review, this role is collectively termed *Affective Therapy*. Nine of these studies (650 participants) were conducted on participants diagnosed with dementia. In total, 11 reported positive findings including reductions in depression scores, agitation scores, and increases in quality of life scores. Whilst these studies were evaluating similar

effects of SAR, their intervention design can broadly be divided into two types: one-on-one interactions with SAR or group interactions with SAR.

Eight studies (657 participants) assessed SAR in one-on-one settings whereas the remaining seven studies (232 participants) had group settings. All of the group setting studies reported positive findings, including reduced agitation and depression levels, and higher expression of positive emotions. Of the 8 one-on-one interaction studies, only 5 report positive findings. Indeed, 2 of these studies(12)(13) report negative findings with increased agitation and worsening dementia, respectively.

These contrasting set of results could indicate a mechanism of how elderly users gain emotional benefit from SAR. A Japanese pilot study(14) assessed group interactions of 26 subjects with Paro and found significant improvements in mood scores during the intervention period. Of note, the authors commented on improved sociability between subjects. As discussed later, several studies(15)(16)(17)(18)(19) demonstrate that SAR can increase the sociability of subjects within groups, which may play a direct role in the mood changes seen here.

Notwithstanding this however, a Dutch crossover study(20) compared two types of one-on-one intervention: therapeutic interventions (Paro introduced at times when subject was distressed) and care support interventions (Paro introduced to facilitate activities of daily living). Only the therapeutic intervention showed a significant improvement in the mood score ($p < 0.01$). This suggests that perhaps while group interventions may be better at generating positive emotions, one-on-one interventions may be appropriate to remedy negative emotions.

Some studies in this set also investigated how SAR compared to soft toys in improving general mood and wellbeing of participants. A large Australian randomised

controlled trial(21) of 415 participants with dementia, compared one-on-one interventions with Paro switched 'on' and Paro switched 'off' (placebo Paro) to identify if Paro's additional social capabilities translated into any positive outcomes. The study found Paro was more effective than usual care in improving pleasure and agitation, but was no different to placebo Paro. Similarly, a Japanese study (8) compared the effect of group interactions with Paro and placebo Paro, and again did not demonstrate any differences between the groups.

These results are mimicked by a Danish randomised controlled trial(13) of 100 subjects, which compared interactions with Paro, a living dog or soft toy cat. The study found intervention type did not affect cognitive state, independence or depression scores and did not affect sleep quality. However, depressive scores improved compared to baseline scores in all groups ($p < 0.05$).

Indeed, only 2 small pilot studies found differences between SAR and soft toys. The first (22) showed subjects engaged more with Paro ($p < 0.05$) and showed more positive emotional expressions with Paro ($p < 0.01$) when compared to a stuffed lion. The second (23) was a study on participants with dementia; it showed that agitation scores were only significantly decreased in a toy cat ($p < 0.05$), whereas NeCoRo (SAR - cat-like robot) only improved scores of pleasure and interest ($p < 0.01$ and $p < 0.05$, respectively).

Cognitive Training

Six studies (344 participants) assessed whether SAR can improve aspects of cognition, such as working memory or executive function, and as such this review has termed this set *Cognitive Training*. This set included 4 studies (239 participants) that assessed elderly subjects with dementia, and 2 studies (105 participants) that assessed

elderly subjects who were cognitively intact. Several robot types have been used in this set including 2 closed loop robots capable of learned responses. This means that whilst broad conclusions surrounding the role of SAR in cognitive training can be made, the evidence for any individual robot system is limited. Five of the six studies (133 participants) concluded with positive findings, although there is a breadth of outcome measures used as surrogate markers for cognitive improvement.

Two studies used cognitive tests, such as Mini-Mental State Examination (MMSE) as the primary outcome measure to assess the impact of SAR interactions. The first was a randomised controlled trial(24) of 34 cognitively healthy subjects in Japan using the Nodding Kabochan as the SAR intervention. Subjects either received the fully functional Nodding Kabochan, or a non-functional Nodding Kabochan (control) for 8 weeks. All interactions were one-on-one with the participant and the SAR in the participants home. Only subjects receiving the functional Nodding Kabochan demonstrated an improved cognitive function score ($p < 0.01$) after the study period. This result contrasts with the conclusion of the previous set, Affective Therapy, where it was difficult to distinguish the positive effects between functional SAR and placebo toys. The distinction here may be that the Nodding Kabochan robot is a communication robot that can talk and sing with the user; a function that a placebo toy is incapable of. The communication itself may be key to this study's findings.

The other study that used cognitive tests as an outcome measure for cognition was a 2-phase block randomised controlled trial(25). This Spanish study involved 101 and 110 subjects with dementia, in the respective phases, and assessed the cognitive effects of group interactions with SAR. In Phase 1, the study compared open-loop system robot, Paro, to closed loop robot, NAO, and a control group, treatment as usual. Compared to control

group, Phase 1 showed a decrease in cognitive function scores in the NAO group only ($p < 0.05$) at follow up. Notably, there were no significant differences between NAO and Paro groups at follow-up. This set of results contrasts with the previous study conducted on cognitively healthy subjects in one-on-one settings. Given different robots systems have been used in the studies, it is difficult to establish which factor is responsible for differing results.

Two studies used neuroimaging modalities as outcome measures of interactions with SAR. The first was a South Korean study(26) that used Magnetic Resonance Imaging (MRI) in a randomised controlled trial of 71 cognitively healthy subjects. The primary outcome measure was change in cortical thickness in brains of participants over the 12 week study period. Subjects were randomised into 3 arms: (1) robot-assisted group training using Silbot and Mero (SAR); (2) traditional intervention training, using computer software; or (3) non-intervention arm - control. The study showed attenuation of cortical thinning on MRI in both intervention groups ($p < 0.05$), and estimated it would take 15.3-months for intervention groups to reach the same level of cortical thinning as controls. This study also used neuropsychiatric tests as a secondary outcome measure. Both intervention groups showed greater improvement in the executive function scores than control group ($p < 0.001$). However, in the general cognitive and visual memory tasks, the traditional intervention group had greater improvement than in the robot group. Indeed, the robot group did not outperform the traditional group on any neuropsychological tests. Both Silbot and Mero are communication robots, like the Nodding Kabochan, which may underpin the improvements in executive function. Nonetheless, the SAR arm did not prove to be any more effective than traditional computer software in either outcome measures for cognitive function.

The other study to use a neuroimaging modality was a Japanese pilot study(27) of 14 subjects with dementia. This study investigated the neuropsychological influence of Paro within an interactive group setting by analysing the electroencephalogram (EEG) recordings. They found an increase in cortical neuronal activity in 7 participants, particularly in participants who liked Paro. It is unclear what the clinical meaning of this finding is, and without a control group, one cannot distinguish the effect of SAR from any other stimulating activity on EEG.

The two final studies used game performance as a surrogate marker for cognitive function in participants with dementia. These were very small studies without control groups. The first(28) included 3 subjects and found that verbal encouragement from SAR (Bandit) improved response time in a game quiz, whilst the second study, with 11 participants, concluded the participants' performance in group ball games and individual card games improved following interactions with SAR (AIBO). Again, the clinical utility of this is unclear, and without objective outcome measures or control groups, there is little that can be learned from these studies.

Social Facilitator

Seven studies (230 participants) assessed the utility of SAR as facilitators for improved sociability between subjects or between subjects and other people. As such, this review has titled this role *Social Facilitator*. All of these studies concluded that the respective SAR intervention improved sociability of participants. Five of these studies (210 participants) were conducted with participants who had been diagnosed with dementia. Four of the studies used Paro as the SAR intervention, and two used AIBO, the robotic dog,

which allowed for a greater degree of comparison between the studies. The final study used Sophie and Jack as the SAR intervention.

Most studies used observed behaviour changes on video recording or via a live assessor during the interaction period. One study(16) used a validated communication scale to assess how group Paro interactions affected sociability. The study concluded that after the 4-week programme, a significant improvement in communication and interaction skills were exhibited by subjects ($p < 0.05$) and an increase in activity participation ($p < 0.05$).

Two studies compared SAR to comparative soft toys/ animals. The first was a crossover study(17) of 23 subjects in the US. Subjects were grouped into sessions with Paro, placebo Paro, or no object. The study concluded that the group with Paro engaged in more social interactions than the group with placebo Paro. This suggests that the sociability effects are associated with SAR itself. The authors note that the novelty around SAR may have contributed to the excitement manifested in increased social engagement. However, as this study was conducted over 4 months, any novelty effects would not likely have been sustained.

The other comparative study was another crossover study(29) in the US, which involved 18 female subjects with dementia. Subjects were divided into sessions with AIBO, a real dog, or no object. The study concluded that although all visit types with AIBO, a dog, or no object, stimulated social interaction by the subject, there were no significant differences in the frequency of social behaviours exhibited by the subjects between visit types.

A similar US pilot study(15) of 7 subjects with dementia was instead conducted in a group setting. Subjects within a group were divided into primary users, those individuals who engaged with Paro at any one time, or non-primary users who were defined as

everyone else in the group. The study showed an increase in social interaction over the 7-week period between primary and non-primary users towards each other and towards staff.

This study's results are reflected in two larger, more recent studies, that also investigate effects of group interactions with SAR on participants with dementia. The first is an Australian study(18) of 139 participants conducted over 5 years with Sophie and Jack. The study reported that social engagement increased over the study period. The second was a Norwegian study(19) with 23 participants, that evaluated the effects of group interactions Paro on those with mild-moderate dementia compared to those with severe dementia. The study found that those with mild-moderate dementia paid more attention to Paro than those with severe dementia. The authors note that SAR interventions may need to be more tailored towards the degree of dementia severity. Another finding was that over the 12-week study period, there was a reported increase in interactions with other subjects, and a decrease in interactions with Paro.

Companionship

Three studies (78 participants) assessed the utility of SAR in overcoming the feeling of loneliness and social isolation in the elderly. These studies are collected into a set this review has titled the *Companionship* role. All 3 of the studies examining SAR in this role showed reductions in loneliness scores. None of these studies were conducted on patients with diagnosed dementia. Two studies used AIBO as the intervention, while the third used Paro.

Only one study assessed this in a one-on-one setting. This was a randomised controlled trial(30) of 38 subjects in the US. Subjects were randomised to have weekly one-on-one sessions with a real dog, AIBO, or no object (control). Subjects in the dog or AIBO group

were significantly less lonely than those in the control group at week 7 ($p < 0.05$ respectively). In both intervention groups, there was a higher attachment score compared to the control group. No significant differences were found between the dog and AIBO groups in the assessment of loneliness, or attachment. This is an important finding that suggests an artificial animal (SAR) can be as effective a companion as a pet.

The other 2 studies were conducted in a group setting. The first study was a pilot study(7) of 11 subjects in Japan using AIBO. Mean loneliness scores after the session were significantly lower than those before the session ($p < 0.05$), although longer term benefits were not established. The second was a larger randomised controlled trial(31) of 34 subjects in New Zealand investigated the effects of Paro on loneliness. Subjects were randomised into a Paro group or a control group that attended normal activities. Subjects in the Paro group had a significantly greater decrease in loneliness score at the 12 week follow-up than the control group ($p < 0.05$). This indicated that sustained effects can be achieved.

The last 2 studies do show promising results, however in the context of the previous set of studies, the decreased sense of loneliness may result from increased sociability in the group setting. Sociability was not measured in either study and therefore may act a confounder.

Physiological Therapy

Two studies (33 participants) investigated the effects of SAR on physiological markers, and as such this review titles this set *Physiological Therapy*. This clinical applicability of this set is less clear, but does raise some questions that future studies may be able to answer. Both of these studies used Paro as the SAR intervention.

The first was a pilot study(32) of 21 subjects in New Zealand and investigated the effect of Paro on blood pressure and heart rate. Subjects had a single 10-minute session with Paro where they were free to interact with the robot. Blood pressure and heart rate was recorded before (T1), immediately after (T2) and 5-minutes after (T3) the 10-minute interaction. Overall, no significant changes in blood pressure or heart rate were demonstrated, however the study decided to exclude 4 residents who did not interact or touch the robot. Subsequently, significant decreases in systolic blood pressure ($p < 0.05$) from T1 to T2 were shown, and such decreases were sustained at T3 measurement. Similarly, significant decreases in diastolic blood pressure ($p < 0.05$) from T1 to T2 were shown, however this decrease was not sustained at T3. Between T1 and T3, heart rate significantly decreased ($p < 0.05$).

In the other study(33) of 12 subjects in Japan, physiological effects of interacting with Paro were investigated. Compared to baseline readings, a significant increase in the ratio of urinary 17-ketosteroid:17-hydroxycorticosteroid ($p < 0.01$), by week 4 of Paro being introduced, was found. The authors suggest this confers an improved physiological reaction to stress. A confounder noted was an increase in social interactions with other residents ($p < 0.05$) by week 4, compared to baseline. It is also not clear from this study if Paro played any role in the increased sociability of residents, however in the context of other studies on the topic, it seems likely.

These two studies do not provide much indication of the clinical use of SAR, however they do give a direction for what future studies could investigate further.

Quantitative Comparison

Several studies reported comparative quantitative data, by using the same or similar assessment scales to others within their role category. The data from these studies have been reproduced from the studies and are compiled in Tables 3-5. As different assessment tools were used across studies, a standardised mean score (0 – 100) was generated to allow comparison across similar assessment tools. Five comparable studies were identified in the Affective Therapy, each using a mood scale to assess either anxiety or depression or both, giving rise to 7 comparable sets of data. Of these, 5 showed significant improvements in the mood scores either in the robot intervention group or in the follow-up score, depending on study design.

Four comparable studies were identified in the Cognitive Training set of studies and of these, 3 showed significant improvements in the cognitive scores. Of note, the 2 phases of the Spanish paper(25) have been listed as two separate sets of data as they are different studies with different interventions and different subject numbers; they both use the same control data however, as seen on Table 4.

Finally, three studies with comparable data were identified in the Companionship set of studies, each of which used validated loneliness scales. All of these studies showed significant improvements in loneliness scores in the robot intervention group or in the follow-up score, depending on study design.

No comparative data was identified in the Social Facilitator or Physiological Therapy groups.

Use of Assistive Robot Technology in Elderly Care

Table 1: Characteristics of Selected Studies

Role	Ref.	Participants	Setting	Intervention/ Study Design	Duration	Measures	Outcome
Affective Therapy	Gustafsson C, 2015 (34)	4 subjects (2 men) aged 82-90, with dementia	Dementia care home, Sweden	Supervised one-on-one interaction with JustoCat. Pilot study.	1 session (Unknown time length)/ week for 7 weeks.	QUALID, CMAI and interview	1. No significant changes observed in scales
	Takayanagi K, 2014 (22)	30 subjects (19 with mild/ moderate dementia + 11 with severe dementia), mean age 84.9 years (mild/ moderate), 87.5 (severe)	Nursing care facility, resident's room, Japan	Supervised one-on-one interaction with Paro and Stuffed Lion. Pilot study.	1 session (~15 mins) for each intervention per subject, separated by 3-6 months	Observed behaviour seen in video-recording	*In both groups:* 1. Subjects talked more frequently to PARO ($p < 0.05$) 2. Showed more positive emotional expressions with PARO ($p < 0.01$) *In Mild/ moderate group only:* 1. Showed more negative emotional expressions with Lion 2. Frequencies of touching and stroking and frequencies of talking to staff member were higher with Lion *In Severe group only:* 1. Showed neutral expression more frequently with Lion
	Bemelmans R, 2015 (20)	71 subjects (14 men) with dementia in 2 groups: therapeutic intervention and care support intervention	Psychogeriatric care institutions, Netherlands	Supervised one-on-one interaction with Paro or no intervention. Paro either served as a therapeutic or a care support tool in two separate phases of the study. Crossover study.	5 sessions (~15 mins)/ month for 2 months; each month of therapy was interspersed with a control month. In the therapeutic arm only, additional sessions were given when patient was in distress.	IPPA and Coop/ Wonca after each interaction	1. Therapeutic-related interventions show an increase of IPPA scores by 2 points ($p < 0.01$) 2. Care support intervention showed no effect
	Joranson N, 2015 (35)(36)	53 subjects (20 men) aged 62-95, with a cognitive impairment (MMSE < 25) or diagnosed dementia	Nursing Home, separate room, Norway	Supervised group interaction with Paro or TAU. Randomised controlled trial.	2 sessions (~30 mins)/ week for 12 weeks	Cognitive status, medication, BARS, Norwegian version of CSDD called CDR, QUALID Assessed before (T0), after (T1) and at 3-month follow-up (T2)	1. Reduction in agitation in Paro vs TAU from T0 – T2 ($p < 0.05$) 2. Reduction in depression in Paro vs TAU from T0 – T2 ($p < 0.05$) 3. In those with severe dementia, quality of life scores did not decrease in Paro group from T0 – T2, whereas they did in control. 4. No such difference was found in mild-moderate dementia group
	Moyle W, 2013 (37)	18 subjects, aged >65, with dementia	Nursing Home, Australia	Supervised group interaction with Paro or reading group. Randomised controlled trial.	3 sessions (~45 mins)/ week for 5 weeks	Modified QoLAD, RAID, AES, GDS, Revised Algase Wandering Scale–Nursing Home version and OERS	1. The Paro group had higher QOL-AD and OERS-Pleasure scores following the intervention 2. The Paro group had reduced OERS-Anxiety and OERS-Sadness scores following intervention
	Wada K, 2003(38), (39),	14 subjects (all female) aged 77-	Health Service Facility, Japan	Free group interaction with Paro. Pilot study.	2 sessions (1 hour)/ week for 1 year (and a 5-year	Face scale, GDS and Nursing comments	1. A tendency to improve depression after 8 weeks 2. Improvement in mood

Use of Assistive Robot Technology in Elderly Care

(40)(41), (42), (43), (44), (45)	98, 1 subject without dementia			follow-up)		3. Patients did not lose their interest in the long-term
Thodberg K, 2015 (13)(46)(47)	100 subjects with a mean age of 85.5	Nursing Homes, Denmark	Supervised one-on-one interaction with Paro, Dog or Toy cat. Randomised controlled trial.	2 sessions (10 mins)/ week for 6 weeks	MMSE, GBS, GDS, CAM, Sleep data and BMI	1. Greater interaction with Paro and Dog compared to toy 2. Cognitive and independence scores worsened over study period in all groups ($p < 0.05$) 3. Depression scores improved over study in all group ($p < 0.05$)
Libin A, 2004 (23)	9 subjects (all female) aged 83-98, with dementia	Nursing Home, US	Supervised one-on-one interaction with NeCoRo and Toy cat. Crossover study.	1 session (10 mins) for each intervention	ABMI, LMBS and observations	1. Both cats maintained participant's interest 2. Significant increase in pleasure ($p < 0.01$) and interest ($p < 0.05$) scores whilst playing with NeCoRo 3. Only the toy cat improved agitation scores ($p < 0.05$)
Wada K, 2002(14), (48), (49)	26 subjects (all female) aged 73-93, some subjects had dementia	Day service centre, Japan	Free group interaction with Paro. Pilot study.	3 sessions (~45 mins)/ week for 5 weeks	Summarised POMS, Burnout scale for nursing staff, nursing staff comments	1. Significant improvement in POMS scores ($p < 0.05$) 2. Positive social and psychological effects
Wada K, 2003(8), (50), (51), (52), (53), (54)	23 subjects (6 men) mean age 85	Health Service Facility, Japan	Free group interaction with Paro or placebo Paro. Randomised controlled trial.	4 sessions (1 hour)/ week for 4 weeks	POMS, Face scale, Urinary tests and Nursing comments	1. Improvement in mood and reduction in depression and dejection levels in both groups 2. Urinary results suggest Paro interaction reduces stress
Valentí Soler M, 2015 (25)	Phase 1: 20 subjects (10 men), mean age 77.9 Phase 2: 17 subjects (8 men), mean age 79 All subjects were diagnosed with dementia	Day Care Centre, Spain	Phase 1: Supervised group therapy (Cognitive and Physical) with Nao. Phase 2: Supervised group therapy (Cognitive and Physical) with Paro. Crossover study.	2 sessions (30-40mins)/ week for 3 months	GLDS, sMMSE, MMSE, NPI and AI	*Phase 1:* 1. Increase in deterioration scores 2. Significant decrease was seen in irritability scores and total NPI scores *Phase 2:* 1. Increase in deterioration scores
Lane GW, 2016 (55)	23 subjects (all men) aged 58 – 97, 19 had been diagnosed with dementia	Veteran Residential Care Facility, US	Supervised one-on-one interaction with Paro. Pilot study.	3 sessions (>5 mins) across 1 year	Behaviour (assessment form designed by authors of study – no formal name) Assessments made before, during and after interaction.	1. Increase in observed positive affective and behavioural indicators (e.g. Bright affect, interacting with others, calm) 2. Decrease in observed negative affective and behavioural indicators (e.g. Anxious, sad, yelling) 3. Those who best responded to Paro were calm and approachable at the before interaction
Moyle W, 2017 (21)	415 subjects (101 men) mean age 85. All subjects were diagnosed with dementia	Long term Care Facilities, Australia	Free one-on-one interaction with Paro switched on, Paro switched off or TAU. Cluster-randomised controlled trial.	3 sessions (15 mins)/ week for 10 weeks	Video observations (at baseline and weeks' 1, 5 10 and 15) and CMAI (at baseline and weeks' 10 and 15)	1. Subjects in Paro switched on group were more verbally and visually engaged compared to Paro switched off group 2. Both Paro switched on and switched off groups had reduced neutral affect compared to TAU group 3. Paro switched on was more effective than TAU at improving pleasure and agitation

Use of Assistive Robot Technology in Elderly Care

	Petersen S, 2017 (56)	61 subjects (14 men) mean age 84.3 years. All subjects were diagnosed with dementia	Dementia Units, US	Supervised group interaction with Paro or other activity (music, physical activity and mental stimulation). Randomised controlled trial.	3 sessions (20 mins)/ week for 20 weeks	RAID, CSDD, GLDS, pulse rate, pulse oximetry, GSR and medication	1. Anxiety scores, depression scores and pulse rate in Paro group all significantly decreased over the study period compared to control group
	Moyle W, 2016 (12)	5 subjects (all female) mean age 84 years. All subjects were diagnosed with dementia	Nursing home, Australia	Supervised one-on-one interaction with CuDDler. Pilot Study.	3 session (30 mins)/ week for 5 weeks	CMAI (before and after each session)	1. Agitation scores increased in 4 of the 5 patients across the 5 week study period
Cognitive Training	Tanaka M, 2012 (24)	34 subjects (all female), aged > 65 years, living alone	Participant's home, Japan	Living with Nodding Kabochan or control robot (same design as Nodding Kabochan, but cannot talk or nod). Randomised controlled trial.	8 weeks	Questionnaires, BMI, Cognitive tests, APG, and blood and saliva samples	1. Cognitive scores (MMSE + components of Cognistat) were improved in Nodding Kabochan group 2. Saliva cortisol level was decreased in Nodding Kabochan group 3. Higher reports of loss of fatigue, enhancement of motivation and healing in Nodding Kabochan group
	Valentí Soler M, 2015 (25)	Phase 1: 101 subjects (13 men), mean age 84.7 Phase 2: 110 subjects (11 men), mean age 84.7 All subjects were diagnosed with dementia	Nursing Home, Spain	Phase 1: Supervised group therapy (Cognitive, Musical and Physical) with Paro or Nao or TAU. Randomised controlled trial. Phase 2: Supervised group therapy (Cognitive, Musical and physical) with Paro or Dog or TAU. Randomised controlled trial.	2 sessions (30-40mins)/ week for 3 months	GLDS, sMMSE, MMSE, NPI, APADEM-NH and the QUALID	*Phase 1:* 1. Decreased apathy in NAO and Paro groups 2. Increased delusions in the NAO group 3. Increased irritability in both robot groups 4. Decrease in scores on the MMSE, but not the sMMSE, in the NAO group 5. There were no significant differences between NAO and Paro groups *Phase 2:* 1. Increase QUALID scores in the Paro group compared to the TAU group 2. Increased hallucinations and irritability in both the Paro and Dog groups compared to the TAU group 3. Increased disinhibition in Paro group compared to Dog group 4. Decreased night-time behaviour disturbances in the Paro group compared to Dog group 1
	Kim GH, 2013 (26)	71 healthy subjects, aged >60, based in community	Assessment centre, South Korea	Supervised group interaction with either Silbot and Mero robots (robot cognitive training) or on-screen quiz (traditional cognitive training) or received no cognitive	5 sessions (90 mins)/ week for 12 weeks	MRI, Neuropsychometric tests and Alzheimer's Disease Assessment Scale	1. An attenuation of cortical thinning in both intervention groups 2. Robot therapy showed significantly reduced cortical thinning in the right and left anterior cingulate cortices and small areas of right inferior temporal cortex compared to traditional intervention 3. Global topological organization of white matter

Use of Assistive Robot Technology in Elderly Care

				training (control). Randomised controlled trial.			corticocortical networks was decreased in the control group and the rate of decrease was significantly less in both the intervention groups 4. Robot therapy had greater nodal strength in the left rectus gyrus 5. The intervention groups showed greater improvement in the executive function 6. In the general cognitive and visual memory tasks, the traditional intervention group had greater improvement than in the robot group 7. The robot group did not outperform the traditional group on any neuropsychological test
	Tapus A, 2009 (28), (57), (58), (59)	3 subjects (all female) aged >70 with dementia (some reports say 4 subjects, with 1 male)	Care facility, US	Individual interaction (musical, cognitive game) with Bandit (compared to an on-screen simulation of Bandit in some reports). Pilot study.	1 session (20 mins)/ week for 12 months	sMMSE, Response time, Correctness evaluation and Questionnaire	1. Robot encouragement improved response time
	Hamada T (60)	11 subjects with dementia	Nursing home, Japan	Interaction with AIBO, either individually playing a card game or in a group playing a ball game. Pilot study.	1 session/ day for 5 days	Frequency of activity in video observation	1. Improvement in game performance
	Wada K, 2014 (27), (61)	14 subjects (4 men) mean age 79.2, with dementia	Clinic, Japan	Free group interaction with Paro. Pilot study.	1 session (20 mins)	EEG recording, Questionnaire	1. Improvement in cortical neurons activity of 7 patients, especially in patients who liked the robot
Social Facilitator	Kramer SC, 2009 (29)	18 subjects (all female) with dementia	Nursing Home, participants room, US	Supervised one-on-one interaction with AIBO, Dog or no object. Crossover study.	1 visit (~ 3 mins)/ week for 3 weeks (each week is a different interaction)	Observed behaviour seen in video-recording	1. All visits generate interactive behaviour with visitor
	Sabanovic S, 2013 (15)	7 subjects with dementia	Dementia rehabilitation wing, US	Supervised group interaction with Paro. Pilot study.	1 session (30-45 mins)/ week for 7 weeks	Observed behaviour of primary and non-primary interactor seen in video-recording	1. PARO increases activity in particular modalities of social interaction, which vary between primary and non-primary interactors 2. PARO improved activity levels
	Sung H-C, 2015 (16)	12 subjects (9 men), mean age 77.25	Residential care facility, Taiwan	Supervised group interaction with Paro. Pilot study.	2 sessions (30 mins)/ week for 4 weeks	ACIS, Activity Participation scale	1. Significant improvement in communication and interaction skills 2. Significant improvement in activity participation
	Kidd CD, 2006 (17), (62)	23 subjects, aged 60-104, with high functioning in one nursing home and schizophrenia	Nursing Homes, US	Supervised group interaction with Paro switched on, Paro switched off or no object. Crossover study.	1 session (20 mins)/ 2 weeks (in site A) or per month (in site B) for 4 months (8 sessions vs 4 sessions)	Questionnaire and Observation	1. In switched on Paro group, there was an increase in social interactions; even more in the presence of caregivers or experimenters 2. Switched on Paro also generated feel-good experiences

Use of Assistive Robot Technology in Elderly Care

		and/ or dementia in the other					
	Sakairi K, 2004 (63)	8 subjects (2 men) aged 68-89, with dementia	Group home, Japan	One-on-one interaction with AIBO. Pilot study.	1 session (30 mins)	N-dementia scale, MMSE, behaviour scale and video observation	1. Improving communication with staff in a group home and establishment of friendly relations with occupants
	Chu M, 2017 (18)	139 subjects (95 men) aged from 65 – 90, with dementia	Residential care facilities, Australia	Supervised group interaction with Sophie and Jack. Observational study.	2 sessions (4-6 hours) across 5 years	Behaviour (assessment form developed by authors –no formal name). Assessments made every 5 minutes during session.	1. Increase in social engagement of subjects across the 5-year study period
	Jøranson N, 2016 (19)	23 subjects (7 men) aged from 62 – 92. All subjects had a dementia diagnosis	Nursing homes, Norway	Supervised group interaction with Paro. Observational study.	2 sessions (30 mins)/ week for 12 weeks	Observed behaviour as seen in video recording	1. Subjects with mild to moderate dementia paid more attention to Paro than those with severe dementia 2. Over the study period there was an increase in interactions with other subjects, and a decrease in interactions with Paro
Companionship	Banks MR, 2008 (30)	38 subjects	Nursing Home, US	Free one-on-one interaction with AIBO/ dog or no object. Randomised controlled trial.	1 session (30 mins)/ week for 8 weeks	Modified LAPS, UCLA LS	1. Dog and AIBO therapy equally reduced loneliness compared to control (more improvement in most lonely participants; in the control group, the most lonely became more lonely) 2. Residents became significantly and equally attached to AIBO and dog. 3. Attachment was not the mechanism for reduced loneliness in dog or AIBO therapy
	Robinson H, 2013 (31)(64)	34 subjects, aged >55 years	Retirement Home, New Zealand	Group or individual interaction with Paro or alternative activity. Randomised controlled trial.	2 sessions (1 hour)/ week for 12 weeks	UCLA LS, GDS, QoLAD, interview questionnaire and observations	1. Loneliness scores significantly decrease in the Paro group compared to control 2. Residents enjoyed sharing, interacting and talking about Paro
	Kanamori M, 2003 (7)	6 subjects (1 man) aged >64 5 separate control subjects used for CgA measurement	Nursing home/ Participant's home, Japan	Free interaction with AIBO. Control group for CgA measurements had no intervention. Pilot study.	4 sessions (1 hour)/ week for 7 weeks	Scores of emotional words, Amount of speech and Satisfaction, AOKLS, SF-36 and salivary CgA	1. Significant reduction of loneliness 2. Improvement in health related quality of life 3. Decrease in salivary CgA, an indicator of sympathetic adrenal system activity 4. Increase in emotional words, amount of speech and satisfaction exhibited
Physiological Therapy	Robinson H, 2015 (32)	21 subjects (7 men) mean age 84.9	Residential care facility, New Zealand	Supervised one-on-one interaction with Paro. Pilot study.	1 session (10 mins)	Blood pressure reading: Before during and after interaction	1. Significant reductions in systolic and diastolic blood pressure 2. Reduced systolic blood pressure was sustained after Paro was taken away 3. Reduced diastolic blood pressure was not sustained after Paro was taken away

Use of Assistive Robot Technology in Elderly Care

						4. Data suggests average heart rate decreased
Wada K, 2008 (33), (65), (66), (67), (68), (69)	12 subjects, aged 67-89, with mixed cognitive function	Residential care facility, Japan	Free individual/ group interaction with Paro. Pilot study.	1 session (9.5 hours)/ day for 4 weeks	Urinary tests, interviews and video recording observation	1. Increase in social interaction and density of social networks 2. Improvement of subjects' vital organs reaction to stress

ABMI, Agitated Behaviours Mapping Instrument; ACIS, Assessment of Communication and Interaction Skills; AES, Apathy Evaluation Scale; AI, Apathy Inventory; AIBO, Artificial Intelligence Robot; AOKLS, Ando Osada and Kodama Loneliness Scale; APADEM-NH, Apathy Scale for Institutionalized Patients with Dementia Nursing Home version; APG, Accelerated Plethysmography; BARS, Brief Agitation Rating Scale; BMI, Body Mass Index; CAM, Confusion Assessment Method; CDR, Clinical Dementia Rating Scale; CgA, Chromogranin A; CMAI, Cohen Mansfield Agitation Inventory; Coop/ Wonca, Mood scale; CSDD, Cornell Scale for Symptoms of Depression in Dementia; GBS, Gottfries-Bråne-Steen scale; GDS, Geriatric Depression Scale; GLDS, Global Deterioration Scale; GSR, Galvanic Skin Response; IPPA, Goal attainment scale; LAPS, Lexington Attachment to Pets Scale; LMBS, Lawton's Modified Behaviour Stream; MMSE, Mini Mental State Examination; NAO, NPI, Neuropsychiatric Inventory; OERS, Observed Emotion Rating Scale; POMS, Profile of Mood States; QoLAD, Quality of Life in Alzheimer's Disease Scale; QUALID, Quality of Life Scale; RAID, Rating Anxiety in Dementia Scale; SF-36, Short Form Health Survey; sMMSE, Severe Mini Mental State Examination; TAU, Treatment As Usual; UCLA LS, University of California Los Angeles Loneliness Scale

Use of Assistive Robot Technology in Elderly Patients

Robot	Description	Number Used In Respective Roles					
		Affective Therapy	Cognitive Training	Social Facilitation	Companionship	Physiological Therapy	Total
AIBO	A non-verbal, dog-like robot with a metallic appearance and the ability of sight, walking and interpreting commands. AIBO can learn, mature and, on human interaction, express emotional responses.	-	1	2	2	-	4
Bandit	A humanoid robot mounted on a wheeled base. Bandit can speak, gesticulate and make facial expressions.	-	1	-	-	-	1
CuDDler	A robotic teddy bear able to move its neck, arms and eyelids. CuDDler moves its limbs and vocally interacts. CuDDler can respond appropriately to the pattern and type of touch.	1	-	-	-	-	1
Jack and Sophie	Sophie and Jack are communication robots that are capable facial recognition, emotion recognition, vocalisation, gestures, emotive expressions, singing and dancing.	-	-	1	-	-	1
JustoCat	A non-verbal, cat-like robot with replaceable fur and similar proportions and weight to a real cat. JustoCat is capable of breathing, purring and meowing and is designed to sit on a persons lap and respond to stroking.	1	-	-	-	-	1
Mero	A humanoid head mounted on a base, capable of head motion, facial expressions and speech.	-	1	-	-	-	1
NAO	A humanoid robot, 58 cm tall, capable of walking, speech, gesticulation and dance. NAO is able to interact with people and can develop new skills and become personalised.	1	1	-	-	-	2
NeCoRo	A non-verbal, cat-like robot designed to move and look like a real cat. NecoRo can interpret its surroundings and move accordingly. NeCoRo can express emotion.	1	-	-	-	-	1
Nodding Kabochan	A small robot, with the appearance of a child-like teddy, that can talk, sing and nod. It is designed to communicate with users. Nodding Kabochan can play exercise and singing games with the user.	-	1	-	-	-	1
Silbot	A penguin-like robot that can speak and detect faces. Silbot can engage with users in conversation, games and provide care through drug regimen reminders.	-	1	-	-	-	1
Paro	A non-verbal, seal-like robot with the ability to move its head and tail, blink and make sounds and has 5 sensory modalities: light, sound, temperature, posture and tactile. Paro will respond to being held or stroked and can learn to respond to its name. Paro has its own rhythms; will at times be playful, and at	9	2	3	1	2	17

other times sleepy and inactive.

Table 2: Description of Socially Assistive Robots used in Included Studies

Use of Assistive Robot Technology in Elderly Patients

Table 3: Data extracted from comparable studies in Affective Therapy studies

AFFECTIVE THERAPY Mood Scores									
			Control			**Intervention**			
Study	Number of subjects	Outcome Scale	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	p-value
Gustafsson C, 2015 (34)	4	CMAI	-	-	-	12.6 (6.3)	13.3 (6.6)	0.7	0.88[a]
Joranson N, 2015 (35)	53	BARS	22 (19)	23.3 (22)	1.3	20.1 (12.8)	13.7 (11.7)	-6.4	0.044[b]
Joranson N, 2015 (35)	53	CSDD	18.2 (12.3)	24.5 (17.3)	6.3	23.7 (12.9)	18.9 (16.8)	-4.8	0.019[b]
Petersen S, 2017 (56)	61	CSDD	-	-	-2.1	-	-	-7.4	0.001[b]
Petersen S, 2017 (56)	61	RAID	-	-	-0.7	-	-	-3.1	0.003[b]
Moyle W, 2013 (37)	18	GDS	-	28.7 (23.3)	-	-	31.3 (19.3)	-	0.72[c]
Thodberg, 2015 (13)	100	GDS	-	-	-	13.3 (6.7; 33.3)*	13.3 (6.7; 23.3)*	-	< 0.05[d]

a – Study compares mean baseline score in intervention group to mean follow-up score in the intervention group

b – Study compares change in mean score from baseline to follow-up in control group to change in mean score from baseline to follow-up in intervention group

c – Study compares mean follow up score of control group to mean follow up score of intervention group

d – Study compares median baseline score in intervention group to median follow-up score in the intervention group

* Median and interquartile range reported

BARS, Brief Agitation Rating Scale; CMAI, Cohen Mansfield Agitation Inventory; CSDD, Cornell Scale for Symptoms of Depression in Dementia; GDS, Geriatric Depression Scale; RAID,

Use of Assistive Robot Technology in Elderly Patients

Table 4: Data extracted from comparable studies in Cognitive Training studies

COGNITIVE TRAINING Cognition Scores									
			Control			**Intervention**			
Study	Number of subjects	Outcome Scale	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score	p-value
Tanaka M, 2012 (24)	34	MMSE	-	-	-	94.0 (5)	99.0 (2.3)	5	< 0.01[a]
Valentí Soler M, 2015 (25) Phase 1	101	MMSE	12.1 (18.1)	10.4 (15.7)	-1.7	11.8 (17.3)	8.1 (15.0)	-3.7	0.022[b]
Valentí Soler M, 2015 (25) Phase 2	110	MMSE	12.1 (18.1)	10.4 (15.7)	-1.7	10.7 (16.5)	9.1 (15.7)	-1.6	0.282[b]
Kim GH, 2013 (26)	71	ADAS-Cog	-	-	-	89.9 (5.1)	92.6 (4.0)	2.7	<0.001[a]

Use of Assistive Robot Technology in Elderly Patients

a – Study compares mean baseline score in intervention group to mean follow-up score in the intervention group

b – Study compares change in mean score from baseline to follow-up in control group to change in mean score from baseline to follow-up in intervention group

ADAS-Cog, Alzheimer's Disease Assessment Scale – Cognitive subscale; MMSE, Mini Mental State Examination

Use of Assistive Robot Technology in Elderly Patients

Table 5: Data extracted from comparable studies in Companionship studies

COMPANIONSHIP Loneliness Scores									
			Control			**Intervention**			p-value
Study	Number of subjects	Outcome Scale	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score (SD)	Mean Baseline Score (SD)	Mean Follow-up Score (SD)	Change in Mean Score (SD)	
Banks MR, 2008 (30)	38	UCLA LS	-	-	5.7 (1.3)	-	-	-6.0 (2.7)	< 0.05[b]
Robinson H, 2013 (31)	34	UCLA LS	-	-	3.8 (10.3)	-	-	-9.0 (12.6)	0.03[b]
Kanamori M, 2003 (7)	5	AOKLS	-	-	-	3.3 (2.2)	1.0 (1.3)	-	<0.05[a]

a – Study compares mean baseline score in intervention group to mean follow-up score in the intervention group

b – Study compares change in mean score from baseline to follow-up in control group to change in mean score from baseline to follow-up in intervention group

AOKLS, Ando Osada and Kodama Loneliness Scale; UCLA LS, University of California Los Angeles Loneliness Scale.

DISCUSSION

The aim of this review is to identify the roles SAR could play in elderly care. Despite the infancy of this field, the qualitative amalgamation of the studies demonstrated 5 roles for SAR.

Evaluation of SAR Technology

This review identifies 5 roles for SAR in elderly care: Affective Therapy, Cognitive Training, Social Facilitation, Companionship, and Physiological Therapy. These roles provide a comprehensive classification of how this technology has been utilised in social and physical care to date.

The first set of studies demonstrated that SAR can be used to improve the overall sense of well-being of users, and alleviate acute states of mood disturbance. Interestingly, interactions conducted in a group setting proved to be more consistently effective than one-on-one interactions. However, a study(20) showed that one-on-one interventions were useful in alleviating states of distress. This result may apply to patients with delirium and future studies are required to explore this possibility. The overall picture suggests that whilst SAR is capable of improving mood of subjects, it does not seem to be much better than a comparative soft toy or placebo robot. This is demonstrated in patient groups with and without dementia.

This was not true for the second set, Cognitive Training, where communication robots were significantly more effective at improving cognitive outcome measures than soft toys. The clearest evidence for SAR in improving cognitive function was found in those who are cognitively healthy. Whilst positive findings have been found in participants with dementia, obscure outcome measures make it difficult to interpret the meaning of the

findings. The South Korean study(26) showed that computer programmes are at least as effective as SAR interventions, and may raise doubts about the cost-effectiveness of using SAR to only improve elderly users cognitive function.

All the studies in the Social Facilitator set demonstrated improved sociability. This is demonstrated in subjects with and without dementia, and across 3 robot systems (AIBO, Paro and Sophie and Jack). When compared in group settings, SAR was shown to be more effective than a comparator, such as a soft toy. In one US study (29), subjects were divided into one-on-one sessions with AIBO, a real dog, or no object at all, and whilst all sessions increased frequency of exhibited social behaviour, the study concluded no significant differences between session type. Conversely, in a different US study (17), participants had group interactions with Paro, placebo Paro, or no object. The study concluded that the group with Paro engaged in more social interactions than the group with placebo Paro. This suggests that the sociability effects are associated with a group setting, and perhaps in the absence of a group of users, these effects may not exist.

The Companionship set all showed positive findings. However, 2 studies were conducted in group settings, and the observed improved loneliness scores may be confounded by the increased sociability seen in aforementioned studies. This set has far fewer studies than the other sets generated in this review, however, the findings are insightful. If animal-like SAR can be as much a companion as a pet, then such technology may have particular utility in care homes, where health and safety concerns regarding pets, such as allergies and infection risks, restrict their use.

The final set, Physiological Therapy, did show positive findings, however are clinically uninterpretable. Nonetheless, these studies create new questions about the use of SAR for future studies to address. For example, one study (32) demonstrated short term reductions

in blood pressure and heart rate following Paro interactions. The potential implications of these results are two-fold: this short-term reduction in cardiovascular markers could reflect results seen in the Affective Therapy set, that show calming effects of Paro. Additionally, it may be the case that these reductions can be sustained for the long-term and that SAR may have a role as a non-pharmacological intervention for hypertension. Future studies may benefit from incorporating blood pressure and heart rate outcome measures, alongside other metrics in longer-term studies.

Whilst the utility of SAR in Affective Therapy or Cognitive Training can be replaced by cheaper, existing alternatives (e.g. soft toys or computer software), the main value of SAR may lie in its multi-domain functionality. This review has identified 5 such domains where a single intervention may be of simultaneous value.

Quality of Selected Studies

Of all 33 included studies, 11 were randomised controlled trials (RCTs), 12 included more than 30 subjects and 16 had a comparative intervention. These metrics are not in their own right indicative of the quality of the studies, however together they do provide a general picture. The quality of studies is not evenly distributed across the set. Of the RCTs, 6 are in the Affective Therapy set, while there are none in the Social Facilitator set. Similarly, 9 studies in the Affective Therapy set have a comparative intervention compared to 2 in the Social Facilitator set.

This review did not exclude studies based on methodology. The rationale is that low-quality studies can offer an insight into the potential utility of SAR, and guide study design improvements for future studies. For example, a Companionship role is a popular concept for SAR among commentators in the literature, but very few studies demonstrating this

have been conducted. Evidence supporting a Companionship role is socially desirable because of its applicability to serve the elderly population. As reported by one of the selected studies(30), AIBO, the robotic dog, was as effective a companion as a real dog. This has real implications for its use, specifically where a real animal companion may be inappropriate.

Although no studies were excluded on the basis of quality, there are several underlying methodological limitations facing the selected studies that need to be addressed. Low quality data complicates the task of establishing clinical applications of SAR. It also risks undermining the field's efforts, or sensationalising exploratory research. Another limitation is the narrow set of robots assessed, primarily Paro. This restricts the applicability of results to wider SAR systems with different functionality.

There is also a concern for cultural bias as around a third of the studies were conducted in Japan alone. Although more recent studies have been conducted in other cultural environments, most notably the US and Australia, it is not clear if the results are universally applicable. Additionally, there is evidence of gender bias. Around two-thirds of the participants were women. This is a concern since men and women as populations have been shown to regard robot technology differently(70), and therefore some of the reported findings may be exaggerated, or diminished by the participant composition.

Another common study design issue relates to the supervision of interactions that are present in 20 of the included studies. Although supervision ensures safety for the user, it risks altering how the participant interacts with the robot and may change how the participant reports the robot's utility; known as the Hawthorn Effect. Whilst this is difficult to control for when the study is not randomised and no comparator is used, direct supervision may lead to subjects reporting greater positive effects than is necessarily the

case. An example where this may be the case is a US study (29) where subjects were divided into supervised sessions with AIBO, a real dog, or no object at all. One would anticipate that sessions with an object (AIBO or a soft toy) would stimulate a greater behavioural response than no object at all. However, the study concluded there were no significant differences between the responses to the sessions, irrespective of whether an object was present or not. This suggests that the positive findings were completely independent of the intervention, and may instead be a consequence of supervision.

Another main limitation of the selected studies is the nature of chosen outcome measures. They are often abstract, with a limited number of studies identifying a direct clinical need or problem. Although around half of the studies included a comparator intervention, it often involved uninspiring activities or no activity at all. This is an unfair comparison and may inflate the value attributed to the results. As momentum grows behind SAR, these study design flaws will need to be addressed if the technology is going to play a clinical role in the future.

Review Limitations

The primary limitation of this review is the validity of the categorisation of studies into the defined roles. The roles were created retrospectively, as part of a discovery process on extracting data from the final set of studies. Whilst they have utility in evaluating the state of the field and providing defined expectations for the technology, they have generalised sets of studies that are very different in quality, design and sometimes outcome. There is also the issue that some studies demonstrated several roles for SAR. The studies were categorised on the basis of the the primary outcome measures, irrespective of whether a secondary outcome measure would fit into another set. A consequence of this is that the

weight of data in the respective roles may be misleading. All outcomes have been reported in Table 1 for purposes of data transparency.

Furthermore, this review has an inadvertent risk of excluding relevant papers in the screening phase. Although high concordance between the reviewers was reported, the large volume of studies that had to be reviewed invites the possibility that relevant publications were excluded. The main reason for the high exclusion rate was because the broad search criteria identified irrelevant robot interventions, such as surgical robots or telecommunication devices. It is unlikely, however, that an additional study would have changed the conclusions of this review.

Finally, the comparison of assessment values between studies illustrated in Tables 3-5, aimed to provide some comparison between studies where different outcome measures were used. The comparison does have limitations, because although each assessment tool was scaled from 0-100, a score of 50 in one measure does not necessarily correlate to 50 in a different scale. This has made it difficult to reach broad conclusions about the sets of studies.

Future of the Field

In order to achieve successful application of SAR in elderly care, future studies should be more conscious of the outcome measure chosen and its translation into care. Some studies used surrogate measures such as frequency of laughter(22), or performance in particular games(60). Whilst these may be desired outcomes, it is not clearly demonstrated how they meet quantifiable needs of the elderly population. It is likely that any application of SAR will incorporate several of the previously defined roles. Therefore larger studies should assess the intervention's impact in the context of these clear roles with validated outcome measures. For example, one study(24) involved a robot staying at home with the elderly

participants for 8 weeks, and assessed its impact using questionnaires, cognitive tests, blood and saliva samples. Whilst the study demonstrated an improvement in cognitive scores and a reduction in saliva cortisol, it did not assess whether living with a robot for 8 weeks had any impact on loneliness. Larger randomised controlled trials using valid comparators are needed to definitively show where SAR is and is not useful in elderly care.

Conclusion

Socially assistive robots have shown potential in elderly care which, in light of recent demographic shifts, promises to reform the delivery of care for the elderly. Although many of the studies described have methodological issues, the size and quality of studies are improving. This review has qualitatively assessed the existing research and comprehensively outlined the state of the field as it stands. In establishing the 5 roles to which SAR can be ascribed, this review intends not to restrict ambition, but to provide a basis for clinical applicability and design of future studies. This review urges that new studies should be clearer about the precise role any robot intervention intends to serve, and use validated measures to assess their effectiveness. Future studies need to demonstrate how SAR can solve real problems in order to shift from novelty to functionality in elderly care.

Acknowledgements

This review would like to thank Peter Forrest for his scientific guidance and participation in discussions for this project, and Sundhiya Mandalia for her support in interpreting of the data and its analysis.

Declaration of competing interests: None

Contribution (Author contributing):

Design of the review (Jordan Abdi, Marcela Vizcaychipi)
Data collection (Jordan Abdi, Ahmed Al-Hindawi)
Data analysis and interpretation (Jordan Abdi, Tiffany Ng, Marcela Vizcaychipi)
Drafting the article (Jordan Abdi, Ahmed Al-Hindawi, Tiffany Ng)
Critical revision of the article (Jordan Abdi, Tiffany Ng, Marcela Vizcaychipi)
Final approval of the version to be published (Marcela Vizcaychipi)

Funding sources: All authors declare independence from funders.

Transparency Declaration: All authors, external and internal, had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. The lead author affirms the manuscript is honest, accurate and that any discrepancies from the study are accounted for.

Data sharing: no additional data available.

Systematic registration number and name of register: NIHR 58672

References

1. United Nations. World Population Prospects. 2015;1. Available from: http://esa.un.org/unpd/wpp/Publications/Files/WPP2015_Volume-I_Comprehensive-Tables.pdf
2. Thomas J. Insights into Loneliness , Older People and Well-being. 2015;1–10. Available from: http://backup.ons.gov.uk/wp-content/uploads/sites/3/2015/10/Insights-into-Loneliness-Older-People-and-Well-being-2015.pdf
3. Victor C, Scambler S, Bond J, Bowling A. Loneliness, social isolation and living alone in later life. Rev Clin Gerontol. 2000;10(4):407–17.
4. Holt-lunstad J, Smith TB, Layton JB. Social Relationships and Mortality Risk : A Meta-analytic Review. PLoS Med. 2010;7(7).
5. Hegel F, Muhl C, Wrede B, Hielscher-Fastabend M, Sagerer G. Understanding Social Robots. Second Int Conf Adv Comput Interact. 2009;(Section II):169--174.
6. Reiser U, Jacobs T, Arbeiter G, Parlitz C. Care-O-bot ® 3 – Vision of a Robot Butler. 2013. 97-116 p.
7. Kanamori M, Suzuki M, Oshiro H. Pilot study on improvement of quality of life among elderly using a pet-type robot. IEEE Int Conf Robot Autom. 2003;107–12.
8. Wada K, Shibata T, Saito T, Tanie K. Effects of Robot Assisted Activity to Elderly People who Stay at a Health Service Facility for the Aged. IEEE/RSJ Int Conf Intell Robot Syst. 2003;(October).
9. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. 2011. Available from: http://handbook.cochrane.org
10. Orwin RG, Vevea JL. Evaluating Coding Decisions. In: The Handbook of Research Synthesis and Meta-Analysis. 1994. p. 177–200.

Use of Assistive Robot Technology in Elderly Patients

11. Moher D, Liberati A, Tetzlaff J, Altman DG, Grp P. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement (Reprinted from Annals of Internal Medicine). PLoS Med [Internet]. 2009;89(9):873–80. Available from: <Go to ISI>://000270260200002

12. Moyle W, Jones C, Sung B, Bramble M. What Effect Does an Animal Robot Called CuDDler Have on the Engagement and Emotional Response of Older People with Dementia ? A Pilot Feasibility Study. Int J Soc Robot [Internet]. Springer Netherlands; 2016;8(1):145–56. Available from: "http://dx.doi.org/10.1007/s12369-015-0326-7

13. Thodberg K, Sorensen LU, Christensen JW, Poulsen PH, Houbak B, Damgaard V, et al. Therapeutic effects of dog visits in nursing homes for the elderly. [Internet]. Psychogeriatrics. United Kingdom: Wiley-Blackwell Publishing Ltd.; 2015. p. 289–97. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc11&NEWS=N&AN=2015-49600-001

14. Wada K, Shibata T, Saito T, Tanie K. Effects of robot assisted activity for elderly people at day service center and analysis of its factors. Proc 4th World Congr Intell Control Autom (Cat No02EX527) [Internet]. 2002;2:1301–5. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1020792

15. Sabanovic S, Bennett CC, Chang W-L, Huber L. PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. IEEE Int Conf Rehabil Robot [Internet]. United States; 2013;2013:6650427. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24187245

16. Sung H-C, Chang S-M, Chin M-Y, Lee W-L. Robot-assisted therapy for improving social interactions and activity participation among institutionalized older adults: a pilot study. Asia

Pac Psychiatry [Internet]. Australia; 2015;7(1):1–6. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24692085

17. Kidd CD, Taggart W, Turkle S. A Sociable Robot to Encourage Social Interaction among the Elderly. IEEE Int Conf Robot Autom. 2006;3972–6.

18. Chu M, Khosla R, Mohammad S, Khaksar S. Service innovation through social robot engagement to improve dementia care quality. 2017;435(September).

19. Jøranson N, Pedersen I, Marie A, Rokstad M, Aamodt G, Olsen C, et al. Group activity with Paro in nursing homes : systematic investigation of behaviors in participants. 2017;(2016):1345–54.

20. Bemelmans R, Gelderblom GJ, Jonker P, de Witte L. Effectiveness of Robot Paro in Intramural Psychogeriatric Care: A Multicenter Quasi-Experimental Study. J Am Med Dir Assoc [Internet]. United States; 2015;16(11):946–50. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=26115817

21. Moyle W, Jones CJ, Mur JE, Thalib L, Beattie ERA, Shum DKH, et al. Use of a Robotic Seal as a Therapeutic Tool to Improve Dementia Symptoms : A Cluster-Randomized Controlled Trial. 2017;18:766–73.

22. Takayanagi K, Kirita T, Shibata T. Comparison of Verbal and Emotional Responses of Elderly People with Mild/Moderate Dementia and Those with Severe Dementia in Responses to Seal Robot, PARO. Front Aging Neurosci [Internet]. Switzerland; 2014;6:257. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=25309434

23. Libin A, Cohen-Mansfield J. Therapeutic robocat for nursing home residents with dementia: preliminary inquiry. Am J Alzheimers Dis Other Demen [Internet]. United States;

2004;19(2):111–6. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med5&NEWS=N&AN=15106392

24. Tanaka M, Ishii A, Yamano E, Ogikubo H, Okazaki M, Kamimura K, et al. Effect of a human-type communication robot on cognitive function in elderly women living alone. Med Sci Monit [Internet]. Poland; 2012;18(9):CR550-7. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=22936190

25. Valentí Soler M, Agüera-Ortiz L, Olazarán Rodríguez J, Mendoza Rebolledo C, Pérez Muñoz A, Rodríguez Pérez I, et al. Social robots in advanced dementia. Front Aging Neurosci [Internet]. Frontiers; 2015 Jan 3 [cited 2016 Mar 14];7:133. Available from: http://journal.frontiersin.org/article/10.3389/fnagi.2015.00133/abstract

26. Kim GH, Jeon S, Im K, Seo SW, Cho H, Noh Y, et al. Structural brain changes after robot-assisted cognitive training in the elderly: A single-blind randomized controlled trial. Alzheimer's Dement [Internet]. 2013 Jul [cited 2016 Mar 14];9(4):P476–7. Available from: http://www.sciencedirect.com/science/article/pii/S1552526013016282

27. Wada K, Shibata T, Musha T, Kimura S. Robot Therapy for Elders Affected by Dementia. 2014 IEEE Long Isl Syst Appl Technol Conf LISAT 2014. 2014;(August).

28. Tapus A. Improving the quality of life of people with dementia through the use of socially assistive robots. Proc - 2009 Adv Technol Enhanc Qual Life, AT-EQUAL 2009. 2009;81–6.

29. Kramer SC, Friedmann E, Bernstein PL. Comparison of the effect of human interaction, animal-assisted therapy, and AIBO-assisted therapy on long-term care residents with dementia. [Internet]. Anthrozoos. Berg Publishers; 2009. p. 43–57. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc6&NEWS=N&AN=2009-

04201-004

30. Banks MR, Willoughby LM, Banks WA. Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. J Am Med Dir Assoc [Internet]. United States; 2008;9(3):173–7. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med5&NEWS=N&AN=18294600

31. Robinson H, Macdonald B, Kerse N, Broadbent E. The psychosocial effects of a companion robot: a randomized controlled trial. J Am Med Dir Assoc [Internet]. United States; 2013;14(9):661–7. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=23545466

32. Robinson H, MacDonald B, Broadbent E. Physiological effects of a companion robot on blood pressure of older people in residential care facility: a pilot study. Australas J Ageing [Internet]. Australia; 2015;34(1):27–32. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24373064

33. Wada K, Shibata T. Social and physiological influences of robot therapy in a care house. [Internet]. Special Issue: Human and robot interactive communication. 2008. p. 258–76. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc6&NEWS=N&AN=2008-08717-006

34. Gustafsson C, Svanberg C, Müllersdorf M. Using a Robotic Cat in Dementia Care. J Gerontol Nurs [Internet]. Senior Lecturer, School of Health, Care, and Social Welfare, Mälardalen University, Eskilstuna and Västerås, Sweden: SLACK Incorporated; 2015 Oct;41(10):46–56 11p. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=110074731&site=ehost-

live

35. Joranson N, Pedersen I, Rokstad AMM, Ihlebaek C. Effects on Symptoms of Agitation and Depression in Persons With Dementia Participating in Robot-Assisted Activity: A Cluster-Randomized Controlled Trial. J Am Med Dir Assoc [Internet]. United States; 2015;16(10):867–73. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=prem&NEWS=N&AN=26096582

36. Jøranson N, Pedersen I, Marie A, Rokstad M, Ihlebæk C. Change in quality of life in older people with dementia participating in Paro-activity : a cluster-randomized controlled trial. 2016;(June).

37. Moyle W, Cooke M, Beattie E, Jones C, Klein B, Cook G, et al. Exploring the effect of companion robots on emotional expression in older adults with dementia: a pilot randomized controlled trial. J Gerontol Nurs [Internet]. United States; 2013;39(5):46–53. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=23506125

38. Wada K, Shibata T, Saito T, Sakamoto K, Tanie K. Psychological and Social Effects of One Year Robot Assisted Activity on Elderly People at a Health Service Facility for the Aged. Proc IEEE Int Conf Robot Autom. 2003;(April):2785–90.

39. Wada KWK, Shibata TST, Saito TST, Sakamoto KSK, Tanie KTK. Robot assisted activity at a health service facility for the aged for 17 months: an interim report of long-term experiment. IEEE Work Adv Robot its Soc Impacts, 2005. 2005;127–32.

40. Wada K, Shibata T, Sakamoto K, Tanie K. Quantitative analysis of utterance of elderly people in long-term robot assisted activity. Proc - IEEE Int Work Robot Hum Interact Commun. 2005;2005:267–72.

41. Wada K, Shibata T, Saito T, Tanie K. Psychological and social effects in long-term experiment

of robot assisted activity to elderly people at a health service facility for the aged. 2004 IEEE/RSJ Int Conf Intell Robot Syst [Internet]. 2004;3:3068–73. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-14044273830&partnerID=40&md5=b85e82f0c643a517d76b6eb7a05fe686

42. Wada K, Shibata T, Kawaguchi Y. Long-term robot therapy in a health service facility for the aged-A case study for 5 years. 2009 IEEE Int Conf Rehabil Robot ICORR 2009 [Internet]. 2009;930–3. Available from: http://www.scopus.com/inward/record.url?eid=2-s2.0-70449380117&partnerID=40&md5=4bba7ff889ec97a3bd7ebd52ab8af615\nhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5209495\nhttp://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5209495

43. Wada K, Shibata T, Sakamoto K, Saito T, Tanie K. Analysis of utterance in long-term robot assisted activity for elderly people. Proceedings, 2005 IEEE/ASME Int Conf Adv Intell Mechantronics [Internet]. 2005;1(April):31–6. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1500961

44. Wada K, Shibata T, Saito T, Tanie K. Robot assisted activity at a health service facility for the aged for ten weeks: an interim report of a long-term experiment. Int IEEE/EMBS Conf Neural Eng NER. 2006;470–3.

45. Wada K, Shibata T, Saito T, Tanie K. Effects of Three Months Robot Assisted Activity to Depression of Elderly People who stay at a Health Service Facility for the Aged. SICE 2004 Annu Conf. 2004;2709–14.

46. Thodberg K, Sørensen LU, Christensen JW, Poulsen PH, Houbak B, Damgaard V, et al. Therapeutic effects of dog visits in nursing homes for the elderly. 2016;289–97.

47. Thodberg K, Sørensen LU, Videbech PB, Poulsen PH, Houbak B, Damgaard V, et al. Behavioral

Responses of Nursing Home Residents to Visits From a Person with a Dog , a Robot Seal or aToy Cat Behavioral Responses of Nursing Home Residents to Visits From a Person with a Dog , a Robot Seal or aToy Cat. 2016;7936(September 2017).

48. Wada K, Shibata T, Saito T, Tanie K. Analysis of factors that bring mental effects to elderly people in robot assisted activity. IEEE/RSJ Int Conf Intell Robot Syst. 2002;2(October):1152–7.

49. Wada K, Shibata T, Saito T, Tanie K. Robot Assisted Activity for Elderly People and Nurses at a Day Service Center. IEEE/RSJ. 2002;(May):1416–21.

50. Wada K, Shibata T, Saito T, Tanie K. Robot assisted activity to elderly at a health service facility for the aged. Int IEEE/EMBS Conf Neural Eng NER. 2003;

51. Saito T, Shibata T, Wada K, Tanie K. Relationship between interaction with the mental commit robot and change of stress reaction of the elderly. Proc IEEE Int Symp Comput Intell Robot Autom CIRA. 2003;1:119–24.

52. Wada K, Shibata T, Saito T, Tanie K. Relationship between familiarity with mental commit robot and psychological effects to elderly people by robot assisted activity. Proc IEEE Int Symp Comput Intell Robot Autom CIRA. 2003;1:113–8.

53. Wada K, Shibata T, Saito T, Tanie K. Psychological and Social Effects of Robot Assisted Activity to Elderly People who stay at a Health Service Facility for the Aged. IEEE Int Conf Robot Autom. 2004;3:14–9.

54. Wada K, Shibata T, Saito T, Tanie K. Effects of robot-assisted activity for elderly people and nurses at a day service center. Proc IEEE. 2004;92(11):1780–8.

55. Lane GW, Noronha D, Rivera A, et al. Effectiveness of a Social Robot , " Paro ," in a VA Long-Term Care Setting. 2016;(August).

56. Petersen S, Houston S, Qin H, Tague C, Studley J, Health W, et al. The Utilization of Robotic Pets in Dementia Care. 2017;55:569–74.

57. Tapus A, Ţăpuş C, Matarić MJ. The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. 2009 IEEE Int Conf Rehabil Robot ICORR 2009. 2009;924–9.

58. Tapus A, Tapus C, Mataric M. Music therapist robot for individuals with cognitive impairments. ACM/IEEE Int Conf Human-Robot Interact. 2009;297–8.

59. Tapus A, Tapus C, Mataric M. Music Therapist Robot: A Solution for Helping People with Cognitive Impairments. 2009;297–8.

60. Hamada T, Okubo H, Inoue K, Maruyama J, Onari H, Kagawa Y, et al. Robot therapy as for recreation for elderly people with dementia - Game recreation using a pet-type robot -. Proc 17th IEEE Int Symp Robot Hum Interact Commun RO-MAN. 2008;174–9.

61. Wada K, Shibatal T, Musha T, Kimura S. Effects of robot therapy for demented patients evaluated by EEG. 2005 IEEE/RSJ Int Conf Intell Robot Syst IROS. 2005;2205–10.

62. Taggart W, Turkle S, Kidd CD. An Interactive Robot in a Nursing Home : Preliminary Remarks Field Setting : Nursing Homes. Towar Soc Mech Android Sci A COGSCI Work. 2005;1–6.

63. Sakairi K. Research of robot-assisted activity for the elderly with senile dementia in a group home. SICE 2004 Annu Conf. 2004;3:2092–4.

64. Robinson H, Broadbent E, MacDonald B. Group sessions with Paro in a nursing home: Structure, observations and interviews. Australas J Ageing. 2015 Jun;

65. Wada K, Shibata T. Social effects of robot therapy in a care house - Change of social network of the residents for two months. Proc - IEEE Int Conf Robot Autom. 2007;(April):1250–5.

66. Wada K. Social and physiological influences of living with seal robots in an elderly care house for two months. Gerontechnology [Internet]. 2008;235–235. Available from: http://www.gerontechnology.info/Journal/pdf.php?a=645

67. Wada K, Shibata T. Robot therapy in a care house - Change of relationship among the residents

and seal robot during a 2-month long study. Proc - IEEE Int Work Robot Hum Interact Commun. 2007;107–12.

68. Wada K, Shibata T. Robot therapy in a care house - Results of case studies. Proc - IEEE Int Work Robot Hum Interact Commun. 2006;(June 2005):581–6.

69. Wada K, Shibata T. Living With Seal Robots—Its Sociopsychological and Physiological Influences on the Elderly at a Care House. Proc - IEEE Int Work Robot Hum Interact Commun. 2007;23(5):972–80.

70. Schermerhorn P, Scheutz M, Crowell CR. Robot social presence and gender: Do females view robots differently than males? 3rd ACM/IEEE Int Conf. 2008;

Use of Assistive Robot Technology in Elderly Patients

Figure Legends:

Figure 1: Schematic flow diagram of the review process

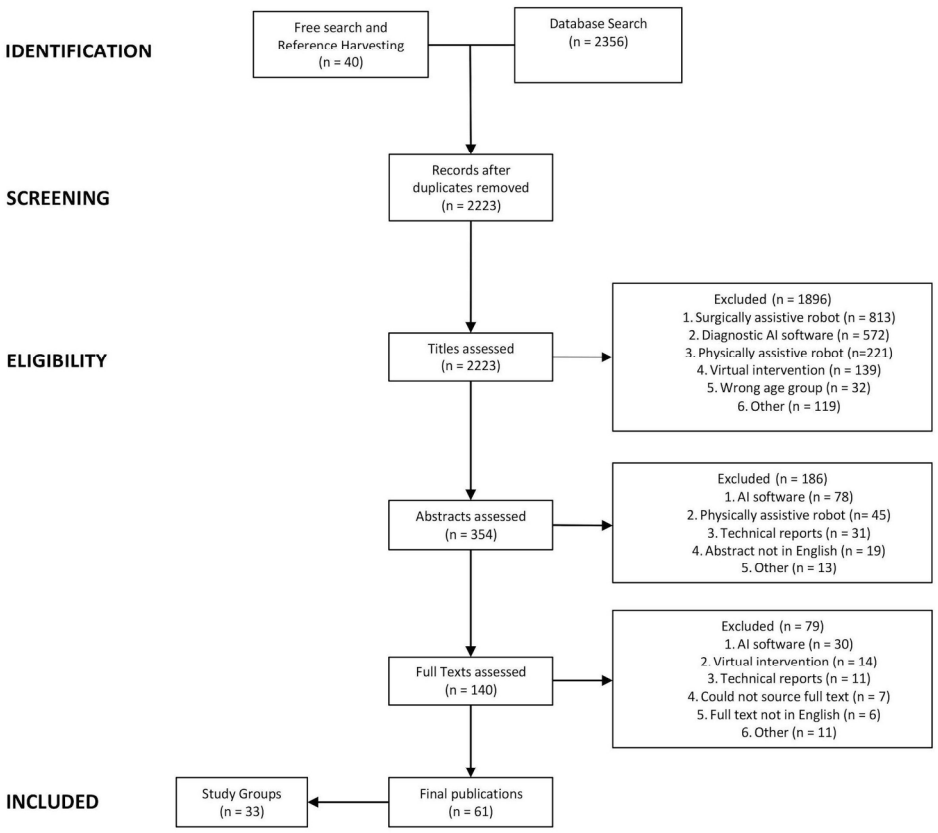

Figure 1: Schematic flow diagram of the review process

179x166mm (300 x 300 DPI)

An example of bibliographical search for PubMEd:

1. Service robot* [Text Word]
2. Therapeutic robot* [Text Word]
3. Socially assistive robot* [Text Word]
4. AIBO [Text Word]
5. Paro [Text Word]
6. Care-o-bot [Text Word]
7. Robotics [MeSH]
8. Artificial Intelligence [MeSH]
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10. Aged, 80 and over [MeSH]
11. Dementia [MeSH]
12. Elder* [Text Word]
13. Senior* [Text Word]
14. Older person [Text Word]
15. Geriatric* [Text Word]
16. Old people [Text Word]
17. 10 or 11 or 12 or 13 or 14 or 15 or 16
18. 9 and 17

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	6
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	6
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	11
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	11
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	11
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	19
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	10
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	19
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	29
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	31
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	34
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	36

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.