JAMDA 18 (2017) 766–773

JAMDA

journal homepage: www.jamda.com

Original Study

# Use of a Robotic Seal as a Therapeutic Tool to Improve Dementia Symptoms: A Cluster-Randomized Controlled Trial

Wendy Moyle PhD [a,b,*], Cindy J. Jones PhD [a,b], Jenny E. Murfield BSc(Hons) [a,b], Lukman Thalib PhD [c], Elizabeth R.A. Beattie PhD [d], David K.H. Shum PhD [a,e], Siobhan T. O'Dwyer PhD [a,f], M. Cindy Mervin PhD [a,g], Brian M. Draper MD [h]

[a] *Menzies Health Institute Queensland, Griffith University, Nathan, Brisbane, Queensland, Australia*
[b] *School of Nursing and Midwifery, Nathan Campus, Griffith University, Nathan, Brisbane, Queensland, Australia*
[c] *Department of Public Health, College of Health Sciences, Qatar University, Qatar*
[d] *School of Nursing, Queensland University of Technology, Kelvin Grove, Brisbane, Queensland, Australia*
[e] *School of Applied Psychology, Mt Gravatt Campus, Griffith University, Brisbane, Queensland, Australia*
[f] *Medical School, University of Exeter, Exeter, United Kingdom*
[g] *Center for Applied Health Economics, School of Medicine, Nathan Campus, Griffith University, Nathan, Brisbane, Queensland, Australia*
[h] *School of Psychiatry, University of New South Wales, Sydney, Australia*

*Keywords:*
Agitation
BPSD
engagement
PARO
mood state
older people

A B S T R A C T

*Objectives:* To test the effects of individual, nonfacilitated sessions with PARO (version 9), when compared against a look-alike plush toy and usual care, on the emotional and behavioral symptoms of dementia for people living in long-term care facilities.
*Design:* Parallel, 3-group, cluster-randomized controlled trial conducted between June 14, 2014, and May 16, 2015.
*Setting:* Twenty-eight long-term care facilities operated by 20 care organizations located in South-East Queensland, Australia.
*Participants:* Four hundred fifteen participants aged ≥60 years, with a documented diagnosis of dementia.
*Intervention:* Stratified by private/not-for-profit status and randomized using a computer-generated sequence, 9 facilities were randomized to the PARO group (individual, nonfacilitated, 15-minute sessions 3 times per week for 10 weeks); 10 to plush toy (same, but given PARO with robotic features disabled); and 9 to usual care. Treatment allocation was masked to assessors.
*Measurements:* Primary outcomes were changes in levels of engagement, mood states, and agitation after a 10-week intervention, assessed by coded video observations (baseline, weeks 1, 5, 10, and 15) and Cohen-Mansfield Agitation Inventory–Short Form (baseline, weeks 10 and 15). Analyses followed intention-to-treat, using repeated measures mixed effects models. Australian New Zealand Clinical Trials Registry (ACTRN12614000508673).
*Results:* Video data showed that participants in the PARO group were more verbally [3.61, 95% confidence interval (CI): 6.40–0.81, $P = .011$] and visually engaged (13.06, 95% CI: 17.05–9.06, $P < .0001$) than participants in plush toy. Both PARO (−3.09, 95% CI: −0.45 to −5.72, $P = .022$) and plush toy (−3.58, 95% CI: −1.26 to −5.91, $P = .002$) had significantly greater reduced neutral affect compared with usual care, whilst PARO was more effective than usual care in improving pleasure (1.12, 95% CI: 1.94–0.29, $P = .008$). Videos showed that PARO was more effective than usual care in improving agitation (3.33, 95% CI: 5.79 −0.86, $P = .008$). When measured using the CMAI-SF, there was no difference between groups.
*Conclusions:* Although more effective than usual care in improving mood states and agitation, PARO was only more effective than a plush toy in encouraging engagement.


This work was supported by the Australian National Health and Medical Research Council [APP1065320].
* Address correspondence to Wendy Moyle, PhD, Menzies Health Institute Queensland, Griffith University, 2.10 Health Sciences (N48), 170 Kessels Rd, Nathan, Brisbane, Queensland, 4111, Australia.
*E-mail address:* w.moyle@griffith.edu.au (W. Moyle).

http://dx.doi.org/10.1016/j.jamda.2017.03.018

Studies from Australia, the United States, and the United Kingdom indicate that at least 50% of residents living in long-term care (LTC) facilities have dementia.[1–3] Of these, over one-half have behavioral and psychological symptoms of dementia (BPSD).[4] These behaviors are often difficult for care staff to manage,[5] and it is common for psychotropic medication to be prescribed as a first-line approach,[6] despite demonstrated adverse effects and inconclusive efficacy.[7] Nonpharmacologic interventions offer an alternate means of managing BPSD, and animal assisted therapies have been successfully used with older people with dementia to ameliorate such symptoms.[8] However, it is not always appropriate for animals to visit LTC facilities (eg, health and safety concerns, residents with a known dislike/fear of animals, and practical issues of looking after an animal), and researchers have sought to investigate how robotic pets may be used instead.

Developed in Japan and modeled on the features of a baby harp seal (Figure 1), PARO is the most common therapeutic pet-type robot used in studies with people with dementia.[9] The therapeutic version (version 9) is an autonomous robot that is similar in weight to a newborn baby, and has 5 sensors that are processed by artificial intelligence software to enable PARO to respond to the user and the environment. Typically active during the daytime, PARO can move its tail and flippers, open and close its eyes, and make sounds similar to a real baby harp seal.

The few randomized controlled trials (RCTs) undertaken to date demonstrate the potential efficacy of using PARO with older people with dementia on measures of anxiety,[10] stress,[10] usage of psychotropic[10–13] and pain medication,[10] depression,[11–13] agitation,[11–13] loneliness,[14] quality of life,[11–13] social interaction,[11–13] and engagement.[11–13] Similarly promising findings on a range of outcomes have also been demonstrated comparing PARO with various control group activities including an interactive reading group,[15] a humanoid robot,[16] a live dog,[16–18] and a stuffed toy.[17–19] Methodological shortcomings limit the reliability and generalizability of these findings, however, and recent editorials and reviews have highlighted the need for more rigorously designed RCTs to further current understanding.[9,20,21]

The aim of this study was to test the effects of individual, nonfacilitated sessions with PARO (version 9), when compared against a look-alike plush toy and usual care, on the emotional and behavioral symptoms of dementia for people living in LTC facilities. We hypothesized that participants in the PARO group would demonstrate improvements in engagement, mood states, and agitation more so than participants in the plush toy and usual care groups.

## Methods

### *Study Design, Setting, and Sample*

This parallel, 3-group, single-blind cluster-RCT was conducted in 28 LTC facilities in South-East Queensland, Australia. A cluster-RCT design was chosen to reduce between-group contamination likely in LTC facilities (ie, inadvertent exposure to activities from different intervention arms because of the nature and layout of facilities); 3 groups enabled PARO to be comparatively assessed against an identical, nonrobotic plush toy and usual care; and the delivery of the interventions in individual, nonfacilitated sessions allowed the unique effect to be evaluated, independent of any extraneous effects of group or facilitator-led sessions. Institutional ethical approval was obtained from Griffith University Human Ethics Committee (NRS/03/14/HREC) and respective care organizations, and approval was obtained from individual facility managers. The study protocol can be read in detail elsewhere.[22]

**Fig. 1.** PARO (version 9) (permission for image given by Dr. Takanori Shibata, National Institute of Advanced Industrial Science and Technology, Japan).

LTC facilities were eligible for inclusion if they were Australian government approved and accredited, provided care to residents with dementia, and were located within a 100-km radius of the Brisbane central business district. Residents were recruited if they were aged ≥60 years and had a documented diagnosis of dementia. Exclusion criteria were respite care admission; dual diagnosis of a serious/persistent mental illness; terminal illness; and unremitting pain/distressing physical symptoms. Potential participants were identified by facility managers, and formally screened against the described criteria by trained research assistants (RAs). Written informed consent was obtained from all participants (if capable) or next-of-kin at the time of enrollment, and participant verbal assent was obtained at each intervention session with PARO or plush toy.

### *Randomization and Masking*

Participating facilities were stratified by private/nonprofit status and randomized in blocks of 3 to PARO, plush toy, or usual care groups. An independent service at Griffith University performed the randomization process, using a computer-generated sequence. Allocation to treatment groups was concealed from facility staff, participants, and families until it was operationally required to begin intervention activities (ie, postbaseline data collection). RAs involved in data collection and data coding were masked to the other intervention groups through assignment of work to 1 group only, and by separate working locations. Intervention RAs were allocated to specific facilities, working with only one of the groups, and were masked to all outcome measurements, as were participants and their families.

### *Procedures*

Participants from facilities allocated to the PARO intervention group received an individual, nonfacilitated, 15-minute session with PARO 3 times per week (Monday, Wednesday, and Friday) for 10 weeks. This duration and frequency of sessions was chosen based on findings from our pilot work.[15] A trained RA gave the PARO to the participant at the start of each session, repeating the same introductory script each time (described elsewhere[22]). RAs left the participant with the PARO to interact with it as they liked, returning after 15 minutes to collect PARO. All sessions were conducted during the afternoon hours of 1:00 PM-5:00 PM (when agitation levels are commonly highest[23]) and wherever the participant was at the time of the allocated session.

Participants in facilities allocated to the plush toy intervention group received the same sessions as described above, but were given a plush toy (PARO with robotic features disabled). Participants in facilities allocated to usual care received care as standard.

### *Outcome Measures*

The 3 primary outcomes of interest were changes in participants' levels of engagement, mood states, and agitation after 10 weeks of the

intervention. Secondary outcomes were short-term effects of the intervention for engagement, mood states, and agitation at weeks 1 and 5, as well as longer-term sustained effects postintervention (week 15) for mood states and agitation.

Video observation of participants was used to measure changes in the outcomes of engagement, mood states, and agitation. The same trained RAs who gave PARO to participants also recorded participants using a small, handheld GoPro Hero video camera, without being intrusive. Each participant was recorded at 5 time points (baseline, weeks 1, 5, and 10, and week 15) for 30 minutes each time. For the PARO and plush toy groups, recordings covered 15 minutes before the intervention and 15 minutes during the intervention. For the usual care group, recordings covered any 30-minute period between 1:00 PM and 5:00 PM. All video data were coded in Noldus Observer XT by trained RAs using the Video Coding Protocol-Incorporating Observed Emotion Scheme,[24] a quantitative measure of the time participants' displayed agitation, affect, and behavioral, verbal, visual, and social engagement during recorded observations. RA inter-rater (kappa = 0.95; $P < .0001$) and intra-rater reliability was very high (kappa = 0.95; $P < .0001$).

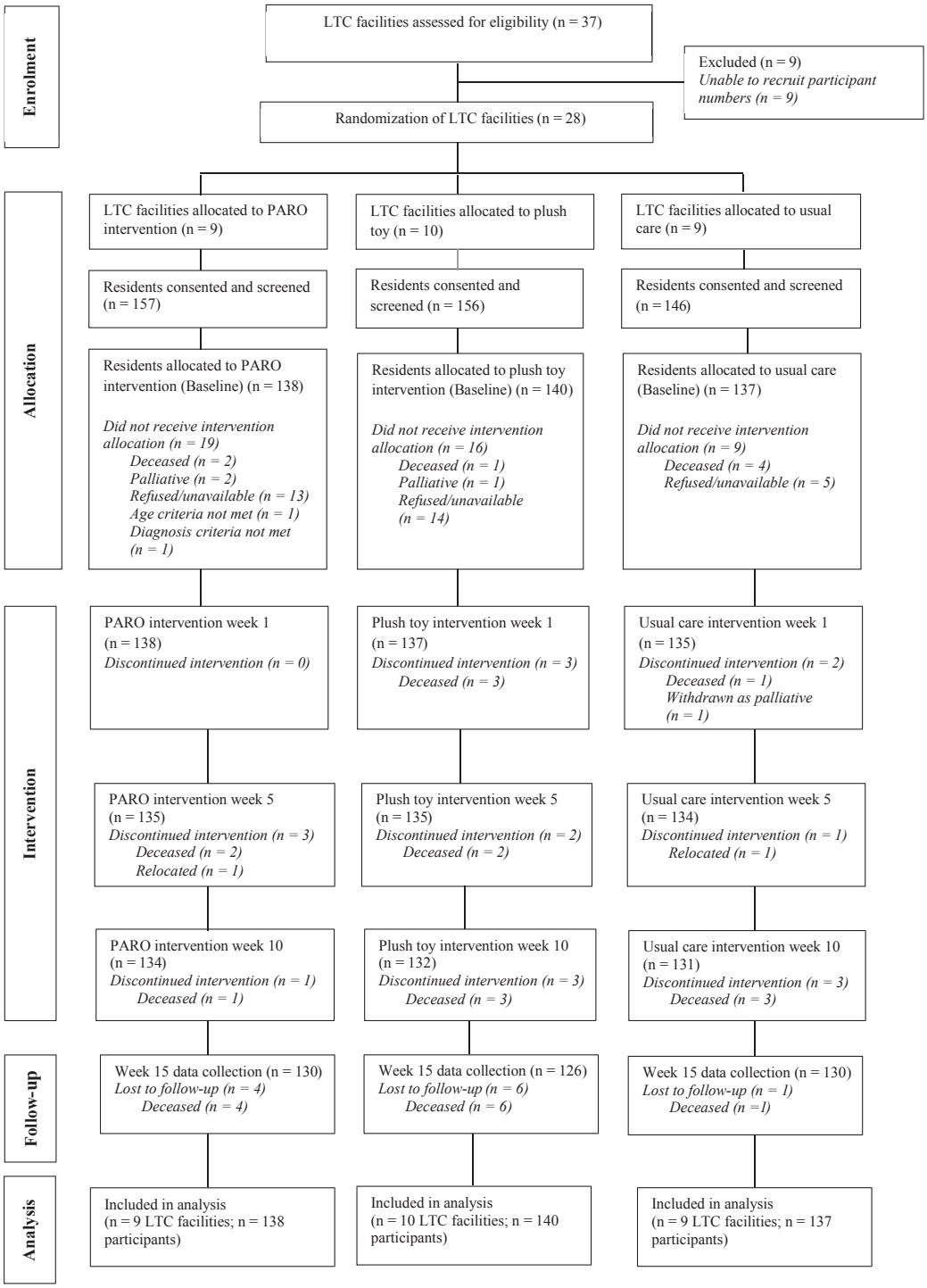


**Fig. 2.** Trial profile.

**Table 1**
Baseline Characteristics

| | PARO | Plush Toy | Usual Care |
|---|---|---|---|
| Number of facilities | (n = 9) | (n = 10) | (n = 9) |
| Number of participants | (n = 138) | (n = 140) | (n = 137) |
| Sex (female) | 101 (73%) | 114 (81%) | 99 (72%) |
| Age (y) | 84 (8.4) | 86 (7.6) | 85 (7.1) |
| RUDAS (total score) | 6.5 (6.5) | 7.1 (6.5) | 8.3 (7.2) |
| CMAI-SF (total score) | 29.0 (10.1) | 30.0 (11.3) | 31.1 (10.3) |
| Taking medication (yes)* | 118 (86%) | 96 (69%) | 103 (75%) |
| Sensory deficit (yes)† | 117/134 (87%) | 125/137 (91%) | 107 (78%) |
| Facility care-type environment:‡ | | | |
| Secure dementia unit | 81 (59%) | 80/138 (58%) | 75 (55%) |
| Facility ward/unit | 57 (41%) | 58/138 (42%) | 62 (45%) |

RUDAS, The Rowland Universal Dementia Assessment Scale: A Multicultural Cognitive Assessment Scale (lower scores indicate greater cognitive impairment); SD, standard deviation.
Data are n (%), mean (SD).
Higher CMAI–SF scores indicate more aggressive/disruptive behavior.
*Includes antidepressants; antipsychotics; anxiolytics and hypnotics; anticonvulsants; analgesics; and anticholinesterase medications.
†Includes hearing; vision; olfaction; touch/pain/tingling; and other deficits.
‡Data not available for all randomized participants.

The Cohen-Mansfield Agitation Inventory–Short Form (CMAI-SF)[25] was also used as a proxy measure of agitation, having established reliability and validity.[26] Facility care staff completed the 14-item CMAI-SF at baseline, week 10, and week 15, rating on a 5-point scale the frequency that participants displayed agitated behaviors during the previous 2-week period. A total summated score is calculated, ranging from 14 to 70, with higher scores indicating greater agitation. At baseline, various cluster- and participant-level information was recorded by RAs.

### *Statistical Analysis*

To compare the effect of the 10-week intervention for our primary and secondary outcomes, a series of repeated measures mixed effects models, using the xtmixed command (adjusted for clustering effects), were undertaken within an intention-to-treat framework. Intraclass correlation coefficients were computed to explore the effect of clustering, and Cohen's d effect sizes were calculated when differences at the alpha level of 0.05 were observed. Video observation data were extracted using Noldus Observer XT (Noldus Information Technology bv, Wageningen, The Netherlands), and all data were analyzed in Stata v 13 (StataCorp, College Station, TX). The trial is registered with the Australian New Zealand Clinical Trials Registry (ACTRN12614000508673). Appendix 1 describes the full details of the analysis undertaken.

## Results

Between June 14, 2014 and May 16, 2015, we enrolled 415 participants from 28 LTC facilities. Of these, 138 participants from 9 facilities were randomly allocated to the PARO intervention, 140 participants from 10 facilities to the plush toy intervention, and 137 participants from 9 facilities to usual care (Figure 2). On average, participants in the PARO group received the intervention 25.8 times out of a possible 30 (95% CI: 24.7–27.0), plush toy 24.1 times (95% CI: 22.7–25.5), and usual care were recorded 29.3 times (95% CI: 28.6–29.9). The profile of facilities and participants included at baseline for the overall sample are described in Table 1. There was similarity between the 3 groups at the beginning of the study.

For our primary outcome of engagement, the PARO group were significantly more verbally (3.61, 95% CI: 6.40–0.81, $P = .011$) and visually engaged (13.06, 95% CI: 17.05–9.06, $P < .0001$) with the object (PARO) than the plush toy group (Table 2). Clinically, the

**Table 2**
Effects of PARO, Plush Toy, and Usual Care on Primary Outcomes after 10 Weeks of the Intervention

| | PARO vs Plush Toy | | | PARO vs Usual Care | | | Plush Toy vs Usual Care | | | ICC |
|---|---|---|---|---|---|---|---|---|---|---|
| | Adj Mean Diff (95% CI) | Effect Size | *P* Value | Adj Mean Diff (95% CI) | Effect Size | *P* Value | Adj Mean Diff (95% CI) | Effect Size | *P* Value | |
| Primary outcomes after 10 wk of intervention | | | | | | | | | | |
| Engagement | | | | | | | | | | |
| Positive behavioral engagement | 6.34 (13.45 to −0.77) | | .080 | | | | | | | 0.000 |
| Using object for social engagement | 1.22 (2.99 to −0.56) | | .179 | | | | | | | 0.048 |
| Positive verbal engagement* | 3.61 (6.40–0.81) | 0.29 | **.011** | | | | | | | 0.009 |
| Visual engagement* | 13.06 (17.05–9.06) | 0.61 | **<.0001** | | | | | | | 0.000 |
| Mood states (affect) | | | | | | | | | | |
| Anger | 0.07 (0.68 to −0.55) | | .826 | 0.40 (0.86 to −0.06) | | .091 | 0.33 (0.78 to −0.12) | | .152 | 0.000 |
| Anxiety/fear | −0.99 (0.04 to −2.02) | | .060 | −0.28 (0.40 to −0.96) | | .422 | 0.71 (1.49 to −0.08) | | .076 | 0.000 |
| Neutral* | 0.50 (2.86 to −1.86) | | .680 | −3.09 (−0.45 to −5.72) | −0.18 | **.022** | −3.58 (−1.26 to −5.91) | −0.20 | **.002** | 0.000 |
| Pleasure* | 0.79 (1.74 to −0.15) | | .100 | 1.12 (1.94 to 0.29) | 0.20 | **.008** | 0.32 (1.10 to −0.45) | | .414 | 0.000 |
| Sadness | 0.48 (1.24 to −0.28) | | .218 | 0.20 (1.00 to −0.60) | | .631 | −0.28 (0.37 to −0.94) | | .400 | 0.000 |
| Agitation | | | | | | | | | | |
| Reduction in agitation* | 1.28 (4.21 to −1.66) | | .393 | 3.33 (5.79–0.86) | 0.12 | **.008** | 2.05 (4.57 to −0.48) | | .112 | 0.000 |
| CMAI–SF | −0.99 (3.78 to −5.76) | | .684 | −1.89 (2.02 to −5.81) | | .343 | −0.90 (4.11 to −5.92) | | .724 | 0.068 |

Note. Bold values are statistically significant ($P < .05$).
Adj mean diff, adjusted mean difference; ICC, intracluster correlation coefficient.
Coding of behavioral, social, verbal, and visual engagement was related to engagement with an object. In the absence of any object in usual care, analyses of significant differences were conducted between PARO and plush toy groups only. For direct video observation data, change score calculations were carried out by subtracting the 15 minutes recorded before the activity from the 15 minutes recorded during the activity. For the CMAI–SF, change scores reflected the difference between the given assessment time point and the values recorded at week 0 baseline. Interpretation of the direction of the adjusted mean difference and effect size depends on the outcome: positive values favor PARO for behavioral, social, verbal and visual engagement, pleasure, and a reduction in agitation; negative values favor PARO for anger, anxiety/fear, neutral, sadness, and CMAI–SF.
Reported effect sizes are Cohen's d, interpreted as 0.2 = small; 0.6 = medium; and 0.8 = large.
*There is a significant overall group effect at the level of $P < .05$.

**Table 3**
Effects of PARO, Plush Toy, and Usual Care on Secondary Outcomes

| Secondary Outcomes | PARO vs Plush Toy | | | PARO vs Usual Care | | | Plush Toy vs Usual Care | | |
|---|---|---|---|---|---|---|---|---|---|
| | Adj Mean Diff (95% CI) | Effect Size | *P* Value | Adj Mean Diff (95% CI) | Effect Size | *P* Value | Adj Mean Diff (95% CI) | Effect Size | *P* Value |
| Engagement | | | | | | | | | |
| Positive behavioral engagement | | | | | | | | | |
| Wk 1 | 5.06 (14.40 to −4.28) | | .288 | | | | | | |
| Wk 5 | 0.10 (10.72 to −10.51) | | .985 | | | | | | |
| Wk 10 | 7.64 (17.30 to −2.02) | | .121 | | | | | | |
| Using object for social engagement | | | | | | | | | |
| Wk 1 | 1.48 (3.58 to −0.63) | | .169 | | | | | | |
| Wk 5 | 0.12 (1.63 to −1.39) | | .879 | | | | | | |
| Wk 10 | 0.59 (1.59 to −0.42) | | .251 | | | | | | |
| Positive verbal engagement | | | | | | | | | |
| Wk 1 | 3.65 (7.58 to −0.28) | | .069 | | | | | | |
| Wk 5 | 2.15 (5.23 to −0.93) | | .171 | | | | | | |
| Wk 10 | 3.60 (5.94 to 1.25) | 0.30 | **.003** | | | | | | |
| Visual engagement | | | | | | | | | |
| Wk 1 | 13.80 (19.92 to 7.69) | 0.59 | **<.0001** | | | | | | |
| Wk 5 | 10.94 (15.65 to 6.24) | 0.60 | **<.0001** | | | | | | |
| Wk 10 | 12.31 (15.54 to 9.09) | 0.68 | **<.0001** | | | | | | |
| Mood states (affect) | | | | | | | | | |
| Anger | | | | | | | | | |
| Baseline | −0.07 (0.37 to −0.50) | | .766 | 0.08 (0.39 to −0.23) | | .597 | 0.15 (0.57 to −0.27) | | .485 |
| Wk 1 | 0.80 (1.71 to −0.11) | | .084 | 1.00 (1.89 to 0.11) | 0.20 | **.028** | 0.19 (0.49 to −0.11) | | .208 |
| Wk 5 | 1.00 (1.99 to 0.01) | 0.22 | **.047** | 1.14 (2.04 to 0.24) | 0.26 | **.013** | 0.14 (0.56 to −0.28) | | .512 |
| Wk 10 | 0.20 (1.38 to −0.98) | | .735 | 0.72 (1.53 to −0.09) | | .082 | 0.51 (1.41 to −0.38) | | .260 |
| Wk 15 | −0.08 (0.14 to −0.30) | | .465 | 0.07 (0.19 to −0.04) | | .220 | 0.16 (0.39 to −0.08) | | .190 |
| Anxiety/fear | | | | | | | | | |
| Baseline | −0.71 (0.62 to −2.04) | | .297 | −0.56 (0.72 to −1.84) | | .392 | 0.15 (0.60 to −0.30) | | .513 |
| Wk 1 | 0.29 (1.18 to −0.61) | | .528 | 0.38 (1.31 to −0.55) | | .423 | 0.09 (0.39 to −0.21) | | .552 |
| Wk 5 | −0.26 (0.64 to −1.16) | | .574 | −0.21 (0.21 to −0.64) | | .323 | 0.04 (0.85 to −0.76) | | .915 |
| Wk 10 | −1.27 (0.18 to −2.72) | | .087 | −0.00 (0.33 to −0.33) | | .997 | 1.27 (2.68 to −0.15) | | .080 |
| Wk 15 | 0.01 (0.78 to −0.76) | | .984 | 0.31 (0.85 to −0.22) | | .250 | 0.30 (0.87 to −0.26) | | .292 |
| Neutral | | | | | | | | | |
| Baseline | 1.59 (6.52 to −3.34) | | .528 | −2.05 (2.24 to −6.34) | | .350 | −3.64 (−0.46 to −6.81) | −0.19 | **.025** |
| Wk 1 | −4.87 (0.65 to −10.39) | | .084 | −6.65 (−0.82 to −12.47) | −0.36 | **.025** | −1.78 (1.96 to −5.51) | | .352 |
| Wk 5 | −2.38 (4.28 to −9.04) | | .483 | −5.46 (−1.85 to −9.06) | −0.32 | **.003** | −3.08 (2.99 to −9.14) | | .320 |
| Wk 10 | −0.59 (3.39 to −4.58) | | .770 | −4.13 (0.48 to −8.73) | | .079 | −3.53 (1.62 to −8.68) | | .179 |
| Wk 15 | −2.50 (1.42 to −6.42) | | .212 | −1.91 (3.00 to −6.82) | | .446 | 0.59 (5.12 to −3.95) | | .800 |
| Pleasure | | | | | | | | | |
| Baseline | 0.40 (1.74 to −0.94) | | .558 | 0.04 (0.93 to −0.84) | | .923 | −0.36 (0.92 to −1.63) | | .582 |
| Wk 1 | 5.51 (8.06 to 2.95) | 0.59 | **<.0001** | 6.62 (9.09 to 4.16) | 0.78 | **<.0001** | 1.12 (2.14 to 0.09) | 0.24 | **.033** |
| Wk 5 | 1.62 (4.66 to −1.43) | | .297 | 4.13 (6.08 to 2.17) | 0.60 | **<.0001** | 2.51 (5.13 to −0.11) | | .060 |
| Wk 10 | 1.20 (3.05 to −0.65) | | .204 | 2.19 (3.91 to 0.46) | 0.35 | **.013** | 0.99 (2.36 to −0.38) | | .156 |
| Wk 15 | 0.70 (1.40 to −0.01) | | .054 | −0.57 (0.37 to −1.50) | | .234 | −1.26 (−0.17 to −2.36) | −0.25 | **.024** |
| Sadness | | | | | | | | | |
| Baseline | −0.20 (0.95 to −1.34) | | .737 | 0.03 (1.21 to −1.14) | | .959 | 0.23 (1.35 to −0.90) | | .692 |
| Wk 1 | −0.73 (0.03 to −1.48) | | .061 | −0.17 (0.68 to −1.02) | | .692 | 0.55 (1.12 to −0.02) | | .057 |
| Wk 5 | 0.35 (1.06 to −0.36) | | .331 | −0.41 (0.58 to −1.40) | | .415 | −0.76 (−0.03 to −1.49) | −0.19 | **.040** |
| Wk 10 | 1.15 (2.55 to −0.24) | | .105 | 0.36 (1.99 to −1.27) | | .664 | −0.79 (0.19 to −1.77) | | .113 |
| Wk 15 | 0.28 (0.59 to −0.03) | | .074 | 0.28 (0.90 to −0.33) | | .364 | 0.00 (0.58 to −0.58) | | .999 |
| Agitation | | | | | | | | | |
| Reduction in agitation | | | | | | | | | |
| Baseline | 1.34 (5.67 to −2.98) | | .543 | 0.60 (5.30 to −4.10) | | .802 | −0.74 (3.62 to −5.11) | | .739 |
| Wk 1 | 1.34 (5.63 to −2.96) | | .542 | 4.56 (8.37 to 0.74) | 0.24 | **.019** | 3.22 (7.46 to −1.02) | | .137 |
| Wk 5 | −2.07 (4.20 to −8.35) | | .518 | 2.02 (6.45 to −2.41) | | .371 | 4.09 (8.93 to −0.74) | | .097 |
| Wk 10 | 1.21 (6.47 to −4.04) | | .651 | 6.05 (10.66 to 1.44) | 0.31 | **.010** | 4.84 (9.93 to −0.25) | | .063 |
| Wk 15 | 0.51 (6.19 to −5.17) | | .861 | −2.10 (3.08 to −7.28) | | .427 | −2.61 (2.03 to −7.24) | | .270 |
| CMAI–SF | | | | | | | | | |
| Wk 10 | −1.19 (3.80 to −6.19) | | .639 | −3.91 (0.46 to −8.27) | | .079 | −2.71 (2.51 to −7.93) | | .308 |
| Wk 15 | −1.84 (3.46 to −7.13) | | .497 | −3.95 (0.23 to −8.13) | | .064 | −2.12 (2.51 to −6.74) | | .370 |

Note. Bold values are statistically significant ($P < .05$).
Adj mean diff, adjusted mean difference.
Coding of behavioral, social, verbal, and visual engagement was related to engagement with an object. In the absence of any object in usual care, analyses of significant differences were conducted between PARO and plush toy groups only. For direct video observation data, change score calculations were carried out by subtracting the 15 minutes recorded before the activity from the 15 minutes recorded during the activity. For the CMAI–SF, change scores reflected the difference between the given assessment time point and the values recorded at week 0 baseline. Interpretation of the direction of the adjusted mean difference and effect size depends on the outcome: positive values favor PARO for behavioral, social, verbal and visual engagement, pleasure, and a reduction in agitation; negative values favor PARO for anger, anxiety/fear, neutral, sadness, and CMAI–SF.
Reported effect sizes are Cohen's d, interpreted as 0.2 = small; 0.6 = medium; and 0.8 = large.

**Table 4**
Mean Change Scores and Standard Deviations for Levels of Engagement, Mood States, and Agitation for PARO, Plush Toy, and Usual Care Groups at Each Assessment Time Point

| | Baseline* | Change in Wk 1* | Change in Wk 5* | Change in Wk 10* | Change in Wk 15* |
|---|---|---|---|---|---|
| | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) | Mean (SD) |
| Engagement | | | | | |
| Positive behavioral engagement | | | | | |
| PARO (n = 137) | | 37.57 (38.96) | 34.18 (40.04) | 30.38 (36.31) | |
| Plush toy (n = 136) | | 32.51 (36.81) | 34.08 (39.38) | 22.76 (34.48) | |
| Using object for social engagement | | | | | |
| PARO (n = 137) | | 2.37 (9.53) | 1.27 (6.33) | 0.70 (5.47) | |
| Plush toy (n = 136) | | 0.97 (6.68) | 1.23 (7.95) | 0.19 (1.27) | |
| Positive verbal engagement | | | | | |
| PARO (n = 137) | | 8.71 (14.63) | 5.75 (13.11) | 6.67 (13.80) | |
| Plush toy (n = 136) | | 5.00 (10.73) | 3.54 (11.83) | 3.01 (10.52) | |
| Visual engagement | | | | | |
| PARO (n = 137) | | 27.90 (26.71) | 21.73 (20.92) | 20.77 (21.74) | |
| Plush toy (n = 136) | | 14.10 (19.78) | 10.78 (15.13) | 8.46 (13.79) | |
| Mood states (affect) | | | | | |
| Anger | | | | | |
| PARO (n = 137) | 0.04 (1.46) | 0.90 (6.97) | 1.11 (6.12) | 0.69 (4.96) | 0.05 (0.51) |
| Plush toy (n = 136) | 0.11 (2.36) | 0.10 (1.46) | 0.11 (2.39) | 0.49 (5.15) | 0.13 (1.31) |
| Usual care (n = 134) | −0.04 (1.19) | −0.09 (1.06) | −0.03 (1.01) | −0.02 (0.84) | −0.02 (1.48) |
| Anxiety/fear | | | | | |
| PARO (n = 137) | −0.52 (7.07) | 0.24 (5.80) | −0.16 (5.70) | −0.07 (1.83) | 0.32 (3.59) |
| Plush toy (n = 136) | 0.19 (7.01) | −0.05 (0.93) | 0.10 (4.99) | 1.19 (9.76) | 0.31 (2.38) |
| Usual care (n = 134) | 0.04 (1.06) | −0.14 (1.69) | 0.06 (0.72) | −0.07 (1.01) | 0.01 (0.68) |
| Neutral | | | | | |
| PARO (n = 137) | −2.41 (18.24) | −5.90 (20.75) | −6.53 (18.93) | −4.43 (16.94) | −2.77 (17.49) |
| Plush toy (n = 136) | −4.00 (19.99) | −1.03 (17.58) | −4.15 (26.25) | −3.84 (18.89) | −0.27 (16.50) |
| Usual care (n = 134) | −0.37 (19.04) | 0.75 (15.30) | −1.08 (15.02) | −0.31 (15.24) | −0.86 (18.03) |
| Pleasure | | | | | |
| PARO (n = 137) | 0.10 (5.00) | 6.15 (11.63) | 4.59 (8.86) | 2.71 (6.99) | 0.07 (2.98) |
| Plush toy (n = 136) | −0.29 (7.54) | 0.65 (6.10) | 2.97 (11.63) | 1.52 (5.91) | −0.62 (4.54) |
| Usual care (n = 134) | 0.06 (4.72) | −0.47 (2.74) | 0.46 (4.14) | 0.53 (5.39) | 0.63 (5.30) |
| Sadness | | | | | |
| PARO (n = 137) | 0.18 (5.35) | −0.63 (5.33) | 0.19 (4.28) | 0.96 (7.99) | 0.22 (4.61) |
| Plush toy (n = 136) | 0.37 (4.26) | 0.09 (1.88) | −0.16 (1.10) | −0.19 (2.22) | −0.07 (0.99) |
| Usual care (n = 134) | 0.15 (4.30) | −0.46 (3.63) | 0.60 (5.47) | 0.60 (6.32) | −0.07 (4.29) |
| Agitation | | | | | |
| Reduction in agitation | | | | | |
| PARO (n = 137) | 0.06 (13.81) | 4.24 (20.87) | 0.32 (24.98) | 4.12 (19.06) | −1.98 (16.20) |
| Plush toy (n = 136) | −1.28 (16.40) | 2.90 (22.15) | 2.39 (22.82) | 2.90 (22.13) | −2.49 (19.24) |
| Usual care (n = 134) | −0.54 (17.19) | −0.32 (17.35) | −1.71 (14.16) | −1.94 (20.27) | 0.12 (18.50) |

| CMAI–SF | Change from Wk 10 to Baseline† | Change from Wk 15 to Baseline† |
|---|---|---|
| | Mean (SD) | Mean (SD) |
| PARO (n = 67) | −2.66 (6.15) | −4.18 (7.67) |
| Plush toy (n = 70) | −1.68 (9.90) | −2.56 (10.44) |
| Usual care (n = 72) | 1.22 (9.24) | −0.25 (9.10) |

Coding of behavioral, social, verbal, and visual engagement was related to engagement with an object. In the absence of any object in usual care, analyses of significant differences were conducted between PARO and plush toy groups only.

*Change score calculations were carried out by subtracting the 15 mins recorded before the activity from the 15 mins recorded during the activity.

†Change scores reflected the difference between the assessment time point and the values recorded at baseline.

differences between the 2 groups were small for verbal engagement (Cohen's d 0.29), yet more pronounced for visual engagement (Cohen's d 0.61).

Regarding mood states, both PARO (−3.09, 95% CI: −0.45 to −5.72, $P = .022$) and plush toy (−3.58, 95% CI: −1.26 to −5.91, $P = .002$) groups demonstrated greater reductions in neutral affect than usual care, whilst PARO (1.12, 95% CI: 1.94–0.29, $P = .008$) was specifically more effective than usual care in improving pleasure (Table 2). The effect sizes for these significant between-group differences were small (Cohen's d, neutral: PARO −0.18, plush toy −0.20; Pleasure: PARO 0.20).

For the third primary outcome of agitation, when measured by observation, PARO (3.33, 95% CI: 5.79–0.86, $P = .008$) demonstrated effectiveness in reducing agitated behaviour compared with usual care, although the size of this difference was small (Cohen's d 0.12) (Table 2). When agitation was measured using the CMAI-SF, there was no significant difference detected between the 3 groups after the 10-week intervention.

In terms of our secondary outcome focused on the short-term effects of the intervention, analyses suggested that there was an initial novelty effect in response to PARO and plush toy. This short-term effect was strongest for visual engagement with the object and for pleasure, with all other significant results found to have clinically small differences between groups (Tables 3 and 4). Specifically, the PARO group were more visually engaged than plush toy at weeks 1 ($P < .0001$) and 5 ($P < .0001$). For mood states, the PARO group showed greater reductions in neutral affect compared with usual care at both weeks 1 ($P = .025$) and 5 ($P = .003$). There were also increases in pleasure at week 1, with PARO more effective than both plush toy ($P < .0001$) and, to a larger extent, usual care ($P < .0001$). This continued to be significant at week 5, with PARO more effective than usual care in improving pleasure ($P < .0001$). At week 5, plush toy reduced sadness more than usual care ($P = .040$). However, in contrast, we also observed increases in levels of anger in the PARO and plush toy groups, and there were differences between PARO and usual care at

weeks 1 ($P = .028$) and 5 ($P = .013$), and plush toy at week 5 ($P = .047$). Finally, the PARO group experienced improved agitation levels at week 1 compared with usual care ($P = .019$).

When looking at the long-term effects at week 15, we found little evidence to suggest sustainability beyond the intervention period (Tables 3 and 4). The only difference showed that the plush toy group experienced a greater reduction in pleasure than usual care ($P = .024$).

## Discussion

In this cluster-RCT, we found partial support for our hypotheses. Our findings showed that participants in the PARO group were more engaged with the object (PARO) than those in the plush toy group, as demonstrated by greater verbal communication and eye contact. Although statistically significant, the observed difference in verbal engagement levels was of a small size clinically. However, for visual engagement, the difference was of a medium magnitude, suggesting a unique and clinically relevant advantage of PARO when its robotic features are enabled. This is important, given that we observed low levels of general engagement for all participants at baseline and postintervention, with participants most likely sitting, walking aimlessly, lying down, or asleep.

Regarding mood states, video data showed that PARO was more effective than usual care in improving pleasure and reducing neutral affect was no different to plush toy. We also observed increases in the levels of anger in the PARO and plush toy groups vs a negligible decrease in usual care at weeks 1 and 5. Review of videos indicated that observed anger was not directed toward the object and was due to other varied reasons for a small group of participants (eg, PARO interrupted their current activity; other residents continually handled PARO; and removal of PARO). Although it is likely that this small group disproportionately skewed the increased levels of anger observed, it is possible that this is a potential adverse effect of PARO.

When looking at participant agitation levels, our findings are less clear. In the videos, PARO was shown to be more effective than usual care in improving agitation levels but was no different to plush toy. However, when measured using the proxy-rated CMAI-SF, there were no differences between any of the 3 groups after 10 weeks. This finding may show that PARO was unable to address all BPSD that were potentially need-driven, dementia-compromised agitation symptoms. Alternatively, it may be that improvements were unable to be detected by observers using the CMAI-SF. We also had data collection difficulties, particularly at week 10, as many care staff failed to complete the measure, and some only partly completed it. To overcome this, we used the conservative method of last observation carried forward, and this may have prevented the detection of group differences.

Our secondary analyses showed short-term effects of PARO and plush-toy at weeks 1 and 5, but little sustainability beyond the intervention period. These findings raise questions about the initial novelty effect of the PARO and plush toy approach and the intervention dose. Participants were given the intervention for short periods of time, based on our pilot work.[15] A more frequent and longer intervention dose may have had a different effect and, in particular, over time.

Our study is the largest and most rigorous of the PARO studies conducted to date. Importantly, we left participants to decide how they would interact and use PARO, whereas all previous studies have used a human to facilitate engagement, making it difficult to determine whether effects are due to the facilitation or the PARO. Our study also used both a plush toy comparison and usual care control, whereas few studies have previously involved a control. Further, the video capture of everyone in the trial, and the use of a rigorous coding measure,[24] allowed for a reliable measurement of the outcomes, which was assured through inter- and intra-reliability testing of RAs. The videos also allowed the research team to review the provision of the protocol by the intervention RAs, thus, ensuring consistency. Lastly, we included participants at all stages of cognitive impairment and the large, representative study population shown by the demographic variables suggest our findings can be generalized (eg, in countries where people with dementia live in similar LTC facilities).

Our study has several limitations. First, the intervention was of a short duration and this might have influenced the findings. However, we undertook all data collection in the same afternoon time period, and the pre-post video recordings enabled an opportunity to see the immediate impact of the PARO and plush toy on the outcome measures. Second, data collection with the CMAI-SF was hindered, with many care staff not completing or partly completing the measure, resulting in the need for a missing data protocol.

## Conclusions

Our findings partly support the efficacy of PARO, but also suggest that, when there are limited resources, a soft toy animal may be used effectively with a person with dementia. However, PARO should not be used to replace staff time, but rather be used during those inevitable periods when staff are otherwise occupied, or when the individual may benefit from comfort from PARO.

## Acknowledgments

We warmly thank all aged-care organizations, LTCs, care staff, residents, and families who so generously took part in the research. Specific thanks are also extended to the following study personnel: Dr Marguerite Bramble (project management), Dr Jasmin Grayson-Collins (project management and cluster leadership), Ms Amanda McNiven (cluster leadership), Dr Billy Sung (data preparation), and Ms Gloria Cheng (data preparation). We also thank Dr Takanori Shibata for the loan of five PAROs.

The sponsor had no role in any aspect of the study design, undertaking, analysis, and interpretation, or in the reporting of findings and preparation of the manuscript.

## Supplementary Data

Supplementary data related to this article can be found online at http://dx.doi.org/10.1016/j.jamda.2017.03.018.

## References

1. Alzheimer's Association. 2016 Alzheimer's disease facts and figures. Alzheimers Dement 2016;12:459–509.
2. Australian Institute of Health and Welfare. Residential aged care in Australia 2010-11: A statistical overview. Cat. no. AGE 68. Canberra: AIHW; 2012.
3. On behalf of Alzheimer's Society, Prince M, Knapp M, et al. Dementia UK: Update. London: Alzheimer's Society; 2014.
4. Australian Institute of Health and Welfare. Dementia in Australia. Cat. no. AGE 70. Canberra: AIHW; 2012.
5. Brodaty H, Draper B, Low LF. Nursing home staff attitudes towards residents with dementia: Strain and satisfaction with work. J Adv Nurs 2003;44:583–590.
6. Seitz DP, Gill SS, Herrmann N, et al. Pharmacological treatments for neuropsychiatric symptoms of dementia in long-term care: A systematic review. Int Psychogeriatr 2013;25:185–203.
7. Livingston G, Kelly L, Lewis-Holmes E, et al. A systematic review of the clinical effectiveness and cost-effectiveness of sensory, psychological and behavioral interventions for managing agitation in older adults with dementia. Health Technol Assess 2014;18:1–226. v–vi.
8. Filan SL, Llewellyn-Jones RH. Animal-assisted therapy for dementia: A review of the literature. Int Psychogeriatr 2006;18:597–611.
9. Bemelmans R, Gelderblom GJ, Jonker P, et al. Socially assistive robots in elderly care: A systematic review into effects and effectiveness. J Am Med Dir Assoc 2012;13:114–120.e1.
10. Petersen S, Houston S, Qin H, et al. The utilization of robotic pets in dementia care. J Alzheimers Dis 2017;55:569–574.
11. Jøranson N, Pedersen I, Rokstad AM, et al. Group activity with Paro in nursing homes: Systematic investigation of behaviors in participants. Int Psychogeriatr 2016;28:1345–1354.
12. Jøranson N, Pedersen I, Rokstad AM, et al. Change in quality of life in older people with dementia participating in PARO-activity: A cluster-randomized controlled trial. J Adv Nurs 2016;72:3020–3033.

13. Jøranson N, Pedersen I, Rokstad AM, et al. 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 2015;16:867–873.
14. Robinson H, Macdonald B, Kerse N, et al. The psychosocial effects of a companion robot: A randomized controlled trial. J Am Med Dir Assoc 2013;14:661–667.
15. Moyle W, Cooke M, Beattie E, et al. Exploring the effect of companion robots on emotional expression in older adults with dementia: a pilot randomized controlled trial. J Gerontol Nurs 2013;39:46–53.
16. Valenti Soler M, Aguera-Ortiz L, Olazaran Rodriguez J, et al. Social robots in advanced dementia. Front Aging Neurosci 2015;7:133.
17. Thodberg K, Sorensen LU, Christensen JW, et al. Therapeutic effects of dog visits in nursing homes for the elderly. Psychogeriatrics 2016;16:289–297.
18. Thodberg K, Sørensen LU, Videbech PB, et al. Behavioral responses of nursing home residents to visits from a person with a dog, a robot seal or a toy cat. Anthrozoos 2016;29:107–121.
19. 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 2014;6:257.
20. Mordoch E, Osterreicher A, Guse L, et al. Use of social commitment robots in the care of elderly people with dementia: A literature review. Maturitas 2013;74:14–20.
21. Burton A. Dolphins, dogs, and robot seals for the treatment of neurological disease. Lancet Neurol 2013;12:851–852.
22. Moyle W, Beattie E, Draper B, et al. Effect of an interactive therapeutic robotic animal on engagement, mood states, agitation and psychotrophic drug use in people with dementia. A cluster randomised controlled trial protocol. BMJ Open 2015;5:e009097.
23. Khachiyants N, Trinkle D, Son SJ, et al. Sundown syndrome in persons with dementia: An update. Psychiatry Investig 2011;8:275–287.
24. Jones C, Sung B, Moyle W. Assessing engagement in people with dementia: A new approach to assessment using video analysis. Arch Psychiatr Nurs 2015; 29:377–382.
25. Werner P, Cohen-Mansfield J, Koroknay V, et al. The impact of a restraint-reduction program on nursing home residents. Geriatr Nurs 1994;15:142–146.
26. Sansoni J, Marosszeky N, Jeon YH, et al. Final report: Dementia outcomes measurement suite project. Centre for Health Service Development. Australia: University of Wollongong; 2007.