

## Original

# Effect and Task of the Dementia Care-assist AI System for Prediction and Prevention of Behavioral and Psychological Symptoms of Dementia

Naoto Okonogi<sup>1,2,4</sup>, Masaharu Hatano<sup>3</sup>, Taiga Fuju<sup>1,2</sup>, Tetsuya Yamagami<sup>4</sup> and Haruyasu Yamaguchi<sup>1</sup>

1 Tokyo Center for Dementia Care Research and Practices, 1-12-1 Takaido-nishi, Suginami, Tokyo 168-0071, Japan

2 Medical Corporation Taiseikai/Uchida Hospital, 345-1 Kuyahara-machi, Numata, Gunma 378-0005, Japan

3 Senior Dementia Institute, 20-14-401 Tsuzuki-ku, Chigasaki-chuo, Yokohama, Kanagawa 224-0032, Japan

4 Gunma University Graduate School of Health Sciences, 3-39-22 Showa-machi, Maebashi, Gunma 371-8511, Japan

## Abstract

**Background & Aims:** In the treatment of people with dementia (PwD), it is necessary to consider behavioral and psychological symptoms of dementia (BPSD). We developed a Dementia Care-assist AI system (DeCaAI) and examined its effectiveness and tasks.

**Methods:** DeCaAI was introduced in 10 dementia group homes. We enrolled 172 PwD and 206 caregivers. The control group received conventional care, and an intervention group received a 19-week intervention. In the intervention group, we fitted vital sensors to PwD and installed sensors in their rooms. Caregivers recorded cases of BPSD. DeCaAI predicted the occurrence of BPSD from the sensor data and care records and notified caregivers of preventive care measures. We examined severity of BPSD, quality of life, and activities of daily living for PwD in both groups. Moreover, we administered a questionnaire to assess the workload of caregivers.

**Results:** Four of 10 facilities were excluded because of insufficient care record data and no final evaluation. There were no significant interactions in any of the evaluations of the PwD and caregivers. Moreover, the questionnaire revealed that improvement of the predictive accuracy and usability of DeCaAI is necessary.

**Conclusion:** We found no significant improvement on BPSD of PwD and workload of caregivers.

## Article Information

### Key words:

Artificial Intelligence,  
behavioral and psychological symptoms of dementia,  
caregivers,  
dementia,  
Internet of Things

### Publication history:

Received: November 13, 2023

Revised: December 11, 2023

Accepted: December 11, 2023

### Corresponding author:

Naoto Okonogi  
Medical Corporation Taiseikai/Uchida Hospital, 345-1  
Kuyahara-machi, Numata, Gunma 378-0005, Japan  
Tel: 0278-23-1231  
E-mail: h202c003@gunma-u.ac.jp

## 1. Introduction

According to the World Alzheimer Report 2015, the number of people with dementia (PwD) worldwide reached 46.8 million in 2015 and it is estimated to increase to 131.5 million by 2050.<sup>1</sup>

In the treatment of PwD, it is necessary to consider the behavioral and psychological symptoms of dementia (BPSD). Kameoka et al. reported that BPSD was significantly correlated with caregivers' burden.<sup>2</sup> In the UK, BPSD is also described as 'challenging behavior' and addressing the 'unmet needs' of PwD is considered important to reduce BPSD.<sup>3</sup> However, when caregivers provide care for BPSD, a highly individualized assessment and response are required; thus, identifying the unmet needs of PwD is often difficult.<sup>4</sup> Moreover, dementia care under the recent spread of the new coronavirus disease 2019 (COVID-19) has increased BPSD and increased caregivers' stress.<sup>5</sup> Therefore, establishing evidence-based dementia care that does not rely on empirical rules or drugs is necessary.

In recent years, information and communication technologies, such as improved robotics, artificial intelligence (AI), and internet of things (IoT) sensor devices, have made remarkable progress. In Japan, the process chart for accelerating the development of AI in the field of health and medical care was released in 2020.<sup>6</sup> However, half of the field-specific research projects related to the development and application of AI are "imaging diagnosis support" field, and there are only a few research

projects on “long-term care and dementia” field in Japan.<sup>7</sup>

Concerning prediction and prevention of BPSD, some preceding research works had shown in the review article of this field.<sup>8</sup> In 2019, Khan et al. investigated predicting agitation in PwD using wearable devices capable of capturing acceleration, blood volume pulses, skin electrical activity, and skin temperature.<sup>9</sup> A pressure mat was placed on the bed to detect heart rate, respiratory rate, and bed exit. In addition, door sensors and video cameras were installed in each room and common areas. The combined sensor data revealed that minute-by-minute changes in the skin’s electrical activity and temperature were effective in predicting daytime agitation.

Initiatives to assess patients’ lives using multiple sensors 24 h a day have been introduced, and in Japan, initiatives to predict BPSD using multiple sensors are being implemented. In 2019, the Ministry of Internal Affairs and Communications started the IoT Creation Support Project ‘Dementia-responsive IoT Service’.<sup>10</sup> The project aimed to predict BPSD by analyzing vital and environmental data from IoT sensor devices. The project found that the likelihood of developing BPSD increased when pulse rate increased, and respiratory rate decreased. However, in addition to the alert, care advice to prevent the predicted BPSD was required. Therefore, the Dementia Care-assist AI system (DeCaAI) was developed. DeCaAI predicts BPSD from IoT sensors and care records data and recommends appropriate care; thus, it may reduce BPSD and contribute to reducing the burden of dementia care.

This study aimed to investigate the effectiveness of using DeCaAI in dementia care facilities. A cluster randomized controlled trial (RCT) design was used to examine the effects on the BPSD, quality of life (QOL) for PwD, and the workload for caregivers.

## 2. Methods

### 2.1 Participants and period

To examine the effect of DeCaAI, 10 dementia group homes were selected. The facilities had at least two units, and one unit was randomly assigned to the intervention and the other to the control group. The facilities, which already equipped IoT sensors, were excluded. The selection of the facilities was based on recommendations from the Japan Association of Dementia Group Homes, and the opportunity method of the research institution. The selection of the facilities was completed in July 2021. The DeCaAI system, including the sensor devices, tablet terminals, audio devices, and Wi-Fi access points, was installed in September 2021. The system operation started on October 4, 2021, and continued for 19 weeks until February 13, 2022.

We enrolled 172 PwD. Inclusion criteria were PwD with the ability to speak. Exclusion criteria were PwD with rank C (bedridden) in the level of independence in daily living for the disabled elderly (daily life independence level), and those on tube feeding.

We performed cluster RCT. The randomization method involved assigning a random number to the two units at each site on the computer, the higher number being the intervention group and the lower number being the control group.

Participants who left during the study period, those with missing assessment items, and those from facilities where BPSD was not recorded for over 7 weeks (one-third of the intervention period) were excluded from the analysis.

A total of 206 caregivers working in the 10 facilities were also included in the study.

### 2.2 DeCaAI system

#### 2.2.1 Development of the DeCaAI system

The schema of DeCaAI is shown in Fig. 1. The information required for BPSD prediction is the PwD’s vital data, room environmental data, and care record

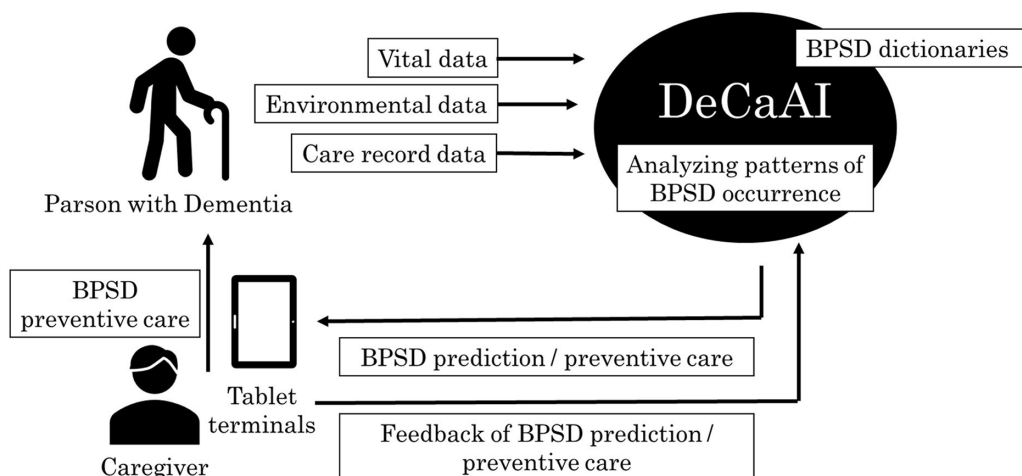


Fig. 1 Scheme of Dementia Care-assist AI system (DeCaAI)

entered by the caregiver. These data are collected to cloud computer via internet communication link. Analyzing these data using AI, DeCaAI predicts BPSD 30 and 60 minutes in advance, and notifies appropriate preventive care to caregivers.

PwD were asked to wear a wristwatch-type vital sensor device (vivosmart4, Garmin) on their wrist. Their pulse rate and step count were measured at 10-minute intervals. At the start, the caregiver explained to the PwD that the device was a substitute for a pedometer and that it should be worn at all times except while bathing and sleeping. Environmental sensors (RS-WFEVS1, RATOC System Corporation) were also installed in the PwD's room and day room (dining room) to measure temperature, air pressure, humidity, illuminance, and carbon dioxide concentration at 10-minute intervals.

For caregivers, three tablet terminals (iPad, Apple Inc.) and three audio devices that were able to transmit audio signals (AirPods, Apple Inc.) with an already installed dedicated application were distributed to each intervention unit. Caregivers recorded BPSD by voice or tap input using tablet terminals and the audio devices. The recording was requested to be performed promptly after BPSD; however, it was also possible to specify the time of occurrence and record it later. When the tablet terminals and audio devices announced BPSD occurrence, appropriate action following DeCaAI-recommended preventive care methods was taken. Other care records were made using the facility's conventional method.

DeCaAI notifies caregivers about a BPSD prediction and recommends care to prevent the predicted BPSD (hereafter referred to as preventive care). The prediction consists of time (30 min, 1 h, 3 h, 6 h, 12 h, 24 h) and type of BPSD. BPSD consists of 17 types (auditory hallucinations, delusions, verbal abuse, assault, wandering, going out without permission, deviant behavior, habitual behavior, excitement, collecting, shouting, depression, apathy, non-response, anxiety, uncleanliness acting, resistance to care, and other). Associations among vital data and environmental data indicators and BPSD onset time were analyzed using a Recurrent Neural Network (RNN), a machine learning technique that predicts future data by learning from accumulated past data. In order to evaluate the accuracy of the created model, a part of the data from the preliminary survey was set aside as "test data" and the rest was used as "training data" to build the model. During preliminary research, this process was

repeated again and again to obtain more appropriate parameter values for improved accuracy of BPSD prediction.

Since the recording time point of vital data and environmental data differs depending on the device, the PI System (OSIsoft, LLC. USA) was used as a preprocessor to align the time sequence of the data. In addition to time sequence analysis, natural language processing is also included in the association of care records and BPSD occurrence times. Therefore, we used a long short-term memory network, which can weigh historical data. BPSD prediction data from the vital and environmental data and those from care records were combined using a Full Connected Layer or RNN to derive the final BPSD prediction results. Trends were derived based on deviations from the mean, not absolute values. An English abstract of the system is available from the Dementia Care Information Network (DCnet) website.<sup>11</sup>

### 2.2.2 BPSD prediction in a preliminary study

As a preliminary study, the validity of the derivation of BPSD predictions was investigated using 54,320 pieces of data from 10 subjects over the past 6 months in a pilot facility (a dementia group home). We identified 3,144 records related to nursing care, of which 595 (18.9%) contained terms related to BPSD. As a result of combining the BPSD types in the care records with vital and environmental data for 24 hours dating back from the time of occurrence, five patterns were derived that predicted BPSD with an accuracy of around 50% (Table 1). It was considered that the number of predictable BPSD types and the prediction accuracy could be increased by having the AI learn from a large amount of data. To confirm the reproducibility of the teacher data, we used 6 months of data from 40 subjects to improve the prediction accuracy. For BPSD prediction using environmental and vital data alone, the percentage of correct predictions within 1 hour and within 12 hours were 63.7% and 63.0%, respectively. On the other hand, for BPSD predictions based on caregiver records alone, the percentage of correct predictions within 1 hour and within 12 hours were 66.2% and 68.8%, respectively. When BPSD prediction was performed by combining environmental and vital data and caregiver records, the highest reproducibility was seen for the combined BPSD prediction: 78.2% of the predictions were correct within 1 hour, and 70.8% were correct within 12 hours.

**Table 1** Occurrence of BPSD and regularity of vital data and environmental data

Occurrence time	Conditions	BPSD	Sampling number	Hit number	Probability
1 Within 3 hours	Temperatures remain high	Auditory hallucinations	20	13	65.0%
2 Within 6 hours	Humidity remains low and illumination remain low	Delusions	14	6	43.8%
3 Within 6 hours	Illumination remains high	Verbal abuse	10	4	40.0%
4 Within 6 hours	Large temperature fluctuations	Wandering	15	10	66.6%
5 Within 12 hours	Temperatures remain high for more than 3 hours	Habitual behavior	10	5	50.0%

BPSD, Behavioral and Psychological Symptoms of Dementia

### 2.2.3 BPSD dictionaries for DeCaAI

To ensure the validity of the preventive care derived by DeCaAI, a corpus of the Dictionary of Dementia-enabling Terms was constructed. The creation of instructional data for the dictionary and the improvement of the corpus' accuracy were carried out by a group of caregivers, nursing experts (one associate professor of nursing, one nurse with 2 years of experience, one visiting caregiver with 10 years of experience, one care support specialist with 25 years of experience), and others.

We asked 10 dementia care experts to develop the Preventive care dictionary. The dictionary, consisting of 227 practical care methods for 25 types of BPSD, was structured using knowledge-structuring AI, and linked to DeCaAI. In case of "wandering", the AI recommends "Let's give them tea and snacks so they can take a break." as preventive care.

Feedback of results to notified prediction and preventive care improved DeCaAI.

## 2.3 Intervention

### 2.3.1 Intervention group

In preparation for the study, we sent the user manual for the DeCaAI application and the procedure manual for device settings to the facilities in advance. Caregivers became familiar with the devices and care record by learning about the video teaching materials in advance. The researchers and research collaborators visited each facility after the start of the intervention period. The settings of the various devices were checked, and the study content was explained to the caregivers in the intervention group.

Caregivers implemented care concerning the recommended preventive care and reflected the results in the care records to improve the prediction by DeCaAI.

### 2.3.2 Control group

Patient care and record keeping were carried out as before. Vital sensors, environmental sensors, and tablet terminals were not used.

## 2.4 Evaluation

Caregivers were asked to collect basic information on PwD and conduct pre-evaluation between September 27, 2021, and October 3, 2021. After a 19-week intervention period from October 4, 2021, to February 13, 2022, a post-evaluation was conducted from February 14, 2022, to February 20, 2022. During the evaluation, the same caregiver conducted both the pre- and post-evaluation for each PwD, without consulting staff members. For PwD evaluation, an input screen was set up in the application of the tablet terminal, and the evaluation results were obtained over the network. The questionnaires for caregivers and forms of basic information on PwD were mailed to each facility, and responses were returned.

### 2.4.1 Evaluation of PwD

Basic information, including age, sex, level of care required, daily life independence level, level of independence in daily living for the elderly with dementia

(dementia independence level), and differential diagnosis of dementia, was recorded. We checked medication for the use of antipsychotic drugs, such as risperidone, quetiapine, and tiapride, and anti-dementia drugs such as donepezil, galantamine, memantine, and rivastigmine.

The primary outcome for PwD was a reduction of the BPSD severity. The Behavioral and Psychological Symptoms of Dementia Questionnaire 25-Item Version (BPSD25Q) was used to assess BPSD severity. Its instructions and a copy of the questionnaire in English can be obtained from DCnet website.<sup>12</sup> The BPSD25Q allows caregivers to judge the severity and burden of 25 BPSD items on a 6-point scale from 0 to 5 (the higher the score, the more severe the BPSD). The scale has three subcategories: hyperactivity (13 items), hypoactivity (six items), and life-related (six items). Its validity and reliability, including correlation with the Neuropsychiatric Inventory-Brief Questionnaire Form, were reported previously.<sup>13</sup>

The secondary outcome for PwD was improvement in QOL. QOL was assessed using the short version of the QOL questionnaire for dementia (short QOL-D). Short QOL-D is calculated from a total of nine items: six in the positive and three in the negative domains. Each item is rated from 1 (not seen) to 4 (often seen). Negative areas are inverted items. Total short QOL-D scores showed a significant negative correlation with the Geriatric Depression Scale score and the apathy score of the Neuropsychiatric Inventory.<sup>14</sup>

The Barthel Index (BI) was used to assess activities of daily living (ADL). BI is a 10-item scale (the higher the score, the higher the level of independence). A high correlation with the Functional Independence Measure, which is widely used as an ADL evaluation tool, was reported.<sup>15</sup>

### 2.4.2 Evaluation of Caregivers

The primary outcome for caregivers was workload. Regarding the workload of caregivers, we evaluated their cumulative work time. The secondary outcome for caregivers was mental health. The World Health Organization-Five Well-Being Index (WHO-5) was used to assess mental health. WHO-5 consists of five questions asking participants about their mood state in the last 2 weeks. Responses are given using a 6-point scale ranging from 'always' to 'never', with scores ranging from 0 to 25 (the higher the score, the better the mental health). It has been reported to have reliability and validity compared to the existing mental health measurement scales, such as the General Health Questionnaire and Philadelphia Geriatric Center Morale Scale, and includes socioeconomic factors and physical factors.<sup>16</sup>

A time study sheet was developed with reference to previous studies to assess the breakdown of caregivers' work content during their working hours.<sup>17</sup> The following items were included: dates, working hours (start to end), time spent recording, time spent dealing with BPSD, and time spent relaxing with residents. 'Dates' and 'working hours' were listed for days worked during the evaluation period week. 'Time spent recording' was



defined as time spent filling out forms and entering data via a tablet and an audio device on a computer. 'Time spent dealing with BPSD' included not only direct care for BPSD but also time spent dealing with indirect issues such as communication and reporting findings. 'Time spent relaxing with residents' was defined as time spent talking, interacting, and working with the residents. For all of these items, the total time was entered in 10-minute increments at the end of the day's work. For staff working in more than one unit, the time was noted for each unit.

The caregivers in the intervention group were required to complete a post-evaluation questionnaire on the effectiveness of the system's use. The questionnaire included the following five items: Q1. Do you think the AI/IoT system is easy to use in dementia care situations? Q2. Do you think AI/IoT systems reduce the time required for care and recording? Q3. Do you think AI/IoT systems accurately predict actual BPSD? Q4. Do you think the BPSD coping strategies proposed by the AI/IoT system are useful for dementia care? Q5. Do you want to continue to use the AI/IoT system in dementia care situations? Answers were rated from a score of 1 (disagree) to 5 (agree). An optional free comment section included system usability, advice, and impressions.

## 2.5 Analysis

IBM SPSS Statistics version 28 was used for statistical analysis. Among the basic information on PwD, 'age' was compared between the intervention and the control groups using an unpaired t-test, whereas were compared using the  $\chi$ -square test or Fisher's direct method.

For evaluating PwD and caregivers, significant differences between the intervention and the control groups in pre-evaluation were analyzed using the Mann-Whitney U-test, and repeated measure analysis of variance was used to examine whether there was an interaction effect.

## 2.6 Ethical considerations

This study was approved by the Ethics Review Committee of the Tokyo Center for Dementia Care Research and Practices (No. 2106).

The study protocol for this cluster RCT has been registered in the UMIN Clinical Trial Registry (UMIN000044715).

Consent for research collaboration was obtained from both caregivers in the target facilities and PwD. Consent was obtained from the caregivers in the facilities through an explanatory document and a consent form. As it was difficult for the researchers to visit the facilities in person to explain the study from the perspective of preventing the spread of COVID-19, a video that explained the study outline and ethical considerations was prepared and distributed to the facilities. Participants and their families were briefed by caregivers using the video and written documents, and consent was obtained from both parties.

The sensor and the care record data obtained in this

study were anonymized and stored via the internet on a server with security measures. The data were only used for this study and stored for 5 years after the end of the study under the responsibility of the Senior Dementia Institute.

## 3. Results

### 3.1 Participants for analysis (Fig. 2)

Of the 172 PwD in the 10 facilities, 156 were randomly allocated and 16 PwD were excluded as shown in Fig. 2.

Of 75 PwD in the intervention group, 29 were excluded due to loss of follow-up. Moreover, 11 from three facilities that had insufficient care records on the system were excluded from the analysis. Thus, in the intervention group, 35 PwD were eligible for analysis. Of the 81 PwD in the control group, 20 were excluded. In addition, 19 PwD from the same three facilities were excluded. In the control group, 42 PwD were eligible for analysis.

Concerning inability to continue care record and evaluation, managers of the excluded facilities indicated the following reasons: infection control of COVID-19; Wi-Fi link problems; and overload of care records.

Of 206 caregivers in the 10 facilities, 192 were randomly allocated and 14 were excluded for leaving the facility before pre-evaluation.

Of 94 in the intervention group, 34 were excluded (two left before the post-evaluation, 25 had missing evaluation data, and seven were from an incomplete evaluation facility). In addition, 14 caregivers from three facilities, where care record were insufficient, were excluded from the analysis. Thus, 46 caregivers from six facilities were eligible for analysis in the intervention group. Of the 98 caregivers in the control group, 37 were excluded (six left before post-evaluation, 22 had missing evaluation data, and nine were from an incomplete evaluation facility). As in the intervention group, 21 caregivers from three facilities, where care record were insufficient, were excluded from the analysis. Thus, 40 caregivers from six facilities were eligible in the control group.

As for the time study sheets, 53 caregivers in the intervention group and 49 in the control group (with some duplicates) were included in the analysis.

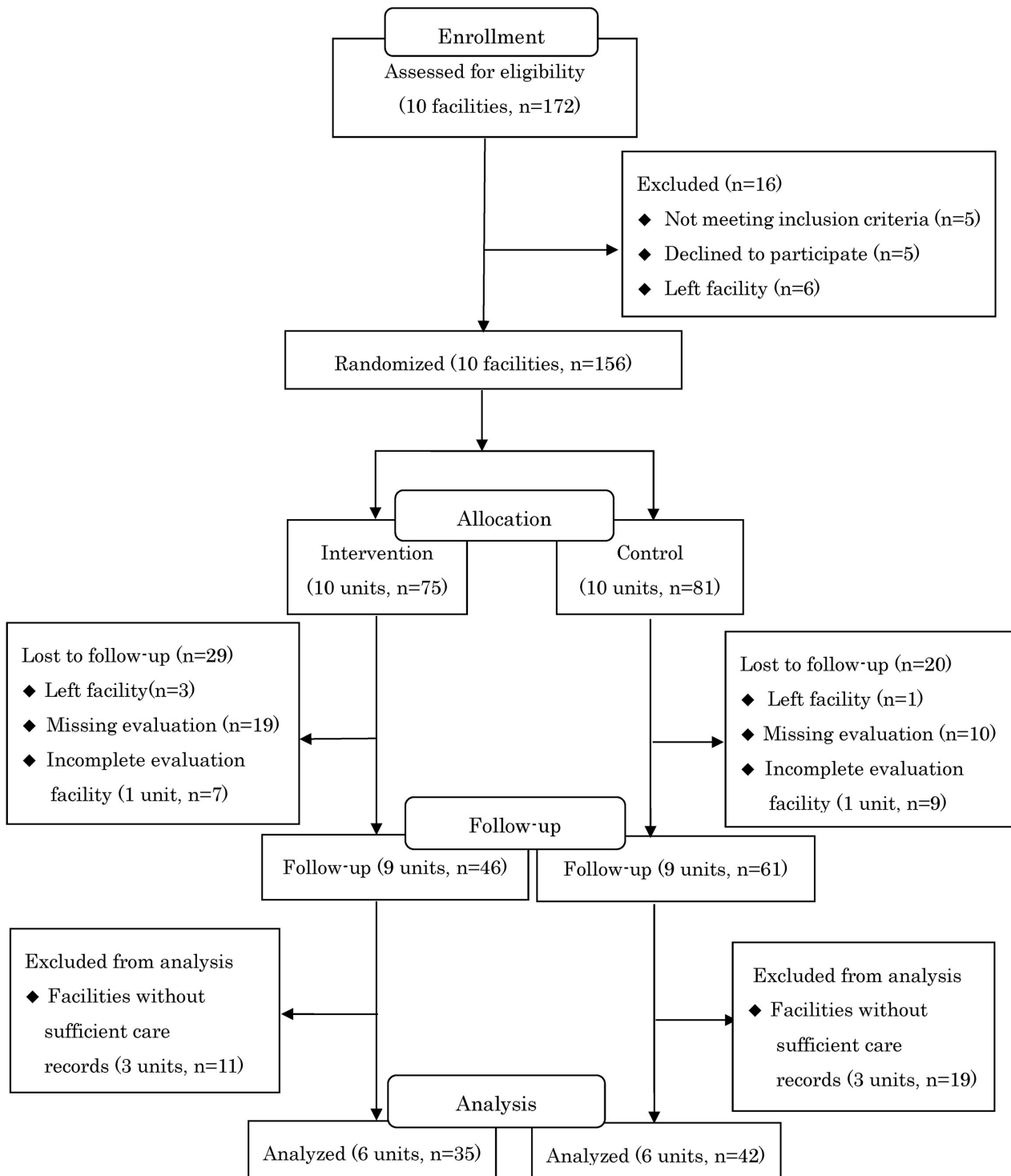
### 3.2 Basic information on PwD

There were no significant differences in any of the following factors between the intervention and the control groups: age, sex, level of care required, daily life independence level, dementia independence level, differential diagnosis of dementia, use of antipsychotic medication, and use of anti-dementia medication (Table 2).

### 3.3 Effectiveness of DeCaAI

#### 3.3.1 Percentage of correct BPSD predictions

The total number of BPSD (17+1 types) recorded by voice input or tap input during the intervention period was 1,512. Total number of BPSD predicted by combin-



**Fig. 2** Flow diagram of the progression through the phases of a cluster randomized trial of two groups

ing vital/environmental data and nursing care records was 417,533, of which 392 BPSD (number of correct predictions) actually occurred. We calculated two markers of the prediction accuracy during the intervention period. Recall (number of correct predictions/number of BPSD occurrences) was 25.9%, and precision (number of correct predictions/number of BPSD predictions) was 0.09%.

### 3.3.2 Evaluation of PwD and caregivers

There were no significant differences among

BPSD25Q, short QOL-D, BI, and WHO-5, and time study of caregivers at pre-evaluation and between the groups (Tables 3 and 4).

### 3.4 The effectiveness questionnaire for caregivers

Questionnaires on the effectiveness of using the system were answered by 71 caregivers in the intervention group. As the questionnaire was conducted anonymously, it was not possible to collate the responses with those who had not yet been evaluated in the pre-and post-evaluations. All respondents were included in the

**Table 2** Basic information of People with Dementia

		Intervention (n=35)	Control (n=42)	P-value
Age (Mean ± SD)		85.1 ± 6.5	85.5 ± 6.9	0.75
Sex	Male	8 (22.9)	3 ( 7.1)	0.10
	Female	27 (77.1)	39 (92.9)	
Level of care required	Support 1	1 ( 2.9)	0 ( 0.0)	0.84
	Support 2	0 ( 0.0)	0 ( 0.0)	
	Care 1	4 (11.4)	8 (19.0)	
	Care 2	11 (31.4)	12 (28.6)	
	Care 3	12 (34.3)	14 (33.3)	
	Care 4	3 ( 8.6)	4 ( 9.5)	
	Care 5	4 (11.4)	4 ( 9.5)	
Daily life independence level	J2	2 ( 5.7)	4 ( 9.5)	0.60
	A1	19 (54.3)	17 (40.5)	
	A2	6 (13.1)	13 (31.0)	
	B1	4 (11.4)	4 ( 9.5)	
	B2	4 (11.4)	4 ( 9.5)	
Dementia independence level	I	0 ( 0.0)	1 ( 2.4)	0.63
	II a	6 (17.1)	4 ( 9.5)	
	II b	5 (14.3)	11 (26.2)	
	III a	16 (45.7)	18 (42.9)	
	III b	5 (14.3)	3 ( 7.1)	
	IV	2 ( 5.7)	4 ( 9.5)	
	M	1 ( 2.9)	1 ( 2.4)	
Diagnosis of dementia	AD	21 (60.0)	24 (57.1)	0.51
	DLB	0 ( 0.0)	3 ( 7.1)	
	FTD	1 ( 2.9)	1 ( 2.4)	
	VD	2 ( 5.7)	4 ( 9.5)	
	No differentiation	11 (31.4)	1 ( 2.4)	
Antipsychotic	Quetiapine	2 ( 5.7)	3 ( 7.1)	0.75
	Tiapride	2 ( 5.7)	3 ( 7.1)	
	Risperidone	2 ( 5.7)	3 ( 7.1)	
	Other	0 ( 0.0)	1 ( 2.4)	
	None	29 (82.9)	32 (76.2)	
Antidementia	Donepezil	6 (17.1)	6 (14.3)	0.48
	Galantamine	1 ( 2.9)	3 ( 7.1)	
	Memantine	7 (20.0)	4 ( 9.5)	
	None	21 (60.0)	29 (69.0)	

AD, Alzheimer's disease dementia; DLB, Dementia with Lewy bodies; FTD, Frontotemporal dementia; VD, Vascular dementia.

**Table 3** Evaluation of People with Dementia

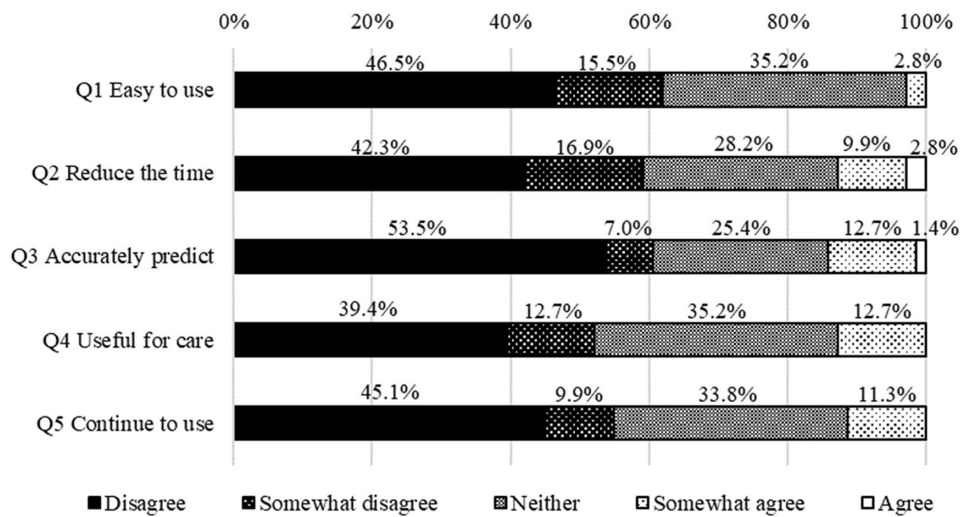
Group		Pre Mean ± SD	Post Mean ± SD	Pre Mann-Whitney (U)	Repeated ANOVA
BPSD25Q severity	Intervention	19.5 ± 14.2	24.3 ± 16.1	$p=0.16$	F=1.20
	Control	14.9 ± 12.2	16.9 ± 12.3		$p=0.28$
BPSD25Q burden	Intervention	14.0 ± 15.4	15.6 ± 17.5	$p=0.24$	F=0.09
	Control	10.3 ± 9.3	11.1 ± 11.5		$p=0.76$
Short QOL-D	Intervention	23.9 ± 5.7	25.3 ± 5.8	$p=0.08$	F=2.72
	Control	26.3 ± 5.9	25.9 ± 6.1		$p=0.10$
BI	Intervention	59.3 ± 28.5	59.4 ± 26.5	$p=0.79$	F=1.23
	Control	60.7 ± 31.0	64.3 ± 29.1		$p=0.27$

BPSD25Q, Behavioral and Psychological Symptoms of Dementia Questionnaire 25-item version; Short QOL-D, short version of the quality of life questionnaire for dementia; BI, Barthel Index.

**Table 4** Evaluation of caregivers

	Group	Pre Mean±SD	Post Mean±SD	Pre Mann-Whitney(U)	Repeated ANOVA
WHO-5	Intervention	11.9± 5.1	12.0± 5.4	p=0.65	F=0.70
	Control	12.5± 4.7	11.8± 5.4		p=0.41
Time study					
Working days/week (days)	Intervention	2.7± 1.5	2.9± 1.3	p=0.99	F=1.31
	Control	2.7± 1.4	2.6± 1.4		p=0.26
Working hours/day (hours)	Intervention	10.1± 3.9	9.4± 3.1	p=0.80	F=1.86
	Control	10.3± 3.3	10.5± 3.6		p=0.18
Time spent recording/day (minutes)	Intervention	40.1±62.4	42.6±64.7	p=0.90	F=0.07
	Control	38.9±28.8	42.8±39.1		p=0.80
Time spent dealing with BPSD/day (minutes)	Intervention	34.4±39.3	36.8±50.5	p=0.16	F=0.10
	Control	48.9±63.6	54.4±74.5		p=0.75
Time spent relaxing with residents/day (minutes)	Intervention	33.2±38.1	30.9±36.2	p=0.61	F=0.04
	Control	37.5±47.7	34.0±43.2		p=0.84

WHO-5, World Health Organisation-Five Well-Being Index; BPSD, Behavioral and Psychological Symptoms of Dementia.

**Fig. 3** Percentage of responses for each item in the questionnaire

analysis. The percentage of responses to each question item is shown in Fig. 3.

The total percentage of 'agree/somewhat agree' responses for Q1, Q2, Q3, Q4, Q5 were 2.8%, 12.7%, 14.1%, 12.7%, and 11.3%, respectively. Furthermore, there were 43 responses to the open-ended questions. The researchers divided the responses into the following three categories: 'prediction' (seven responses), 'system' (four responses), and 'record' (32 responses). Excerpts of free comment are shown in Table 5. As 'prediction' opinions indicated inaccuracies in predicting BPSD. Other opinions were related to the difficulty of using the equipment. 'System' issues included log-in glitches, system errors, and incompatibility with the existing care record system. On the other hand, there were positive opinions in 'record'. These responses were used to improve usability of the system.

## 4. Discussion

### 4.1 Verification of the effectiveness of DeCaAI

Both in PwD and caregivers, the evaluation scales showed no significant interactions between the intervention and control groups. Our findings suggest that care with DeCaAI did not significantly improve BPSD, QOL, and ADL of PwD. Moreover, it did not reduce the workload of caregivers.

The failure to show a significant ameliorative effect may be caused by low accuracy of BPSD prediction by the DeCaAI. Low usability resulted in inadequate reflection of care results to AI for learning. Therefore, recommended preventive care was sometimes inappropriate for the PwD's physical functioning or time of day.

Additionally, due to the impact of COVID-19 pandemic, caregivers may not have had enough time to cooperate fully with the intervention. In the original plan, we intended to conduct frequent visits to facilities to guide the participants on how to use DeCaAI. How-



**Table 5** Comments of the questionnaire, which were used to improve usability of the system.

Categories	Comments
Prediction	<p>Prediction of wandering for PwD who cannot walk, is inappropriate.</p> <p>Inappropriate recommend of BPSD preventive care, such as doing recreational activities even in the middle of the night.</p> <p>Many PwD have difficulty wearing the vital sensors, so the work of checking them regularly is burdensome.</p> <p>I think there are many problems such as system errors.</p>
System	<p>There are problems such as being forced back to the home screen and having to enter the ID and password again to log in.</p> <p>System changes during the COVID-19 pandemic are a burden for caregivers.</p> <p>It is difficult to use because of the separate input fields.</p> <p>Real-time input is not possible.</p>
Record	<p>The voice recognition is not accurate enough and manual recording is faster.</p> <p>Once we get used to the input, we realize that the time required for recording has been reduced.</p>

ever, during the preparation period from July 2021 to September 2021, a state of emergency due to COVID-19 pandemic was declared in Japan.<sup>18</sup> Hence, we were asked to refrain from unnecessary travel, which resulted in insufficient visits to educate caregivers on how to use DeCaAI. Coco et al. reported that education is crucial for accepting technology and helping caregivers to believe in innovation.<sup>19</sup> Moreover, they emphasized the necessity of including caregivers in the process of developing new technology to diminish fears and negative attitudes.

#### 4.2 Task of the DeCaAI in clinical trials

Contrary to the pilot study, the empirical study showed low values for both recall and precision, as valid sample sizes could not be collected for some of the BPSD types. As factors in the failure to collect valid sample sizes, the questionnaire revealed the following:

- 1) It was difficult for PwD to keep wearing wrist-watch-type vital sensors;
- 2) The network environment was bad;
- 3) The usability of voice input and the system was low.
- 4) Excess load for duplicate care record for conventional system and for DeCaAI.

Thus, sufficient data for accurate prediction may not have been collected in this study.

Although the basic PwD information was input to DeCaAI beforehand, the questionnaire showed that BPSD predictions and preventive care proposals were not suitable for individuals. To improve the prediction and prevention of BPSD, it may be necessary to collect more detailed information, such as life history, hobbies, and preferences, as well as mobility and cognitive function information.

Further studies are needed to improve BPSD prediction accuracy and improve system usability. An improved version of DeCaAI is needed to reduce care burden.

## 5. Conclusions

A cluster RCT study of DeCaAI for predicting BPSD and suggesting preventive care was conducted in a dementia group home. However, due to the influence of COVID-19 pandemic and low accuracy of BPSD prediction, neither PwD nor the caregivers were able to show any significant changes in the evaluated scores for 19 weeks. Hence, our findings suggest that the system needs to be improved further for use by frontline caregivers.

## Acknowledgements

We would like to thank all the caregivers and residents of the dementia group home for their cooperation in carrying out this research. We would also like to express our gratitude to the following people for their support: Prof. Akira Utsumi and Shinji Yokoyama of the University of Electro-Communications; Prof. Takuichi Nishimura of the Japan Advanced Institute of Science and Technology; and Prof. Shogo Kojima of the International University of Health and Welfare; Prof. Noriko Shimasue of Saitama Prefectural University; Dr. Tatsuya Kakigi of University of Hyogo; Prof. Mio Ito, Prof. Kunihiko Hayashi, and Prof. Yuki Ideno of Gunma University; and Prof. Kazuko Obayashi of Nihon Fukushi University.

## Disclosure Statement

This study was supported by a grant from Japan Agency for Medical Research and Development under Grant Number JP22us042001. The funding source had no role in the design, practice, and analysis of this study. The authors have no other potential conflicts of interest to disclose.

## References

1. Alzheimer's Disease International. World Alzheimer Report 2015 The Global Impact of Dementia, 2015.
2. Kameoka N, Sumitani S, Ohmori T. Behavioral and psychological symptoms of dementia (BPSD) and care burden: Examination in the facility staff for elderly residents. *J Med Invest* 2020; 67: 236-239.
3. Michelet M, Selbaek G, Strand BH, et al. Associations between unmet needs for daytime activities and company and scores on the Neuropsychiatric Inventory-Questionnaire in people with dementia: a longitudinal study. *Aging Ment Health* 2022; 26: 725-734.
4. Bessey LJ, Walaszek A. Management of Behavioral and Psychological Symptoms of Dementia. *Curr Psychiatry Rep* 2019; 21: 66.
5. Whiting D, Atee M, Morris T, et al. Effect of COVID-19 on BPSD severity and caregiver distress: Trend data from national dementia-specific behavior support programs in Australia. *Alzheimers Dement* 2021; 17: e058454.
6. Ministry of Health, Labour and Welfare. [Hoken iryo bunya AI kaihatsu kasoku consortium. Giron no seiri to kongo no hokosei (Heisei gan nen 6 gatsu 28 nichi sakutei) o fumaeta koteihyo ni tsuite. [cited November 25, 2023.] Available from: <https://www.mhlw.go.jp/content/10600000/000641325.pdf>
7. Takemura S, The challenge to develop and implement artificial intelligence (AI) technologies in health and medical care in Japan. *J Natl Inst Public Health* 2023; 72: 2-13.
8. Yamagami T, Yamaguchi H. Review of BPSD prevention/prediction system using information and communication technology. *Tokyo J Dementia Care Res* 2021; 5: 8-15.
9. Khan SS, Spasojevic S, Nogas J, et al. 41st Agitation detection in people living with dementia using multimodal sensors. *Annu Int Conf IEEE Eng Med Biol Soc* 2019; 3588-3591.
10. Senior Dementia Institute. Dementia-responsive IoT service with using AI. [cited 12 January 2023]. Available from: <https://www.mhlw.go.jp/content/10601000/000468143.pdf>.
11. Tokyo Center for Dementia Care Research and Practices. Abstract of DeCaAI (Dementia Care-assist AI). [cited 12 January 2023]. Available from: [https://www.dcnet.gr.jp/pdf/journal/decaai\\_en\\_20230104.pdf](https://www.dcnet.gr.jp/pdf/journal/decaai_en_20230104.pdf).
12. Tokyo Center for Dementia Care Research and Practices. Instruction of the Behavioral and Psychological Symptoms of Dementia plus Questionnaire (BPSD+Q) and its 25 items version (BPSD25Q). [cited 12 January 2023]. Available from: [https://www.dcnet.gr.jp/pdf/journal/BPSD\\_Q\\_eng230109.pdf](https://www.dcnet.gr.jp/pdf/journal/BPSD_Q_eng230109.pdf).
13. Naito N, Fujii T, Takiguchi Y, et al. Development and evaluation of a new BPSD questionnaire, BPSD+Q. *Tokyo J Dementia Care Res* 2018; 2: 133-145.
14. Terada S, Oshima E, Ikeda C, et al. Development and evaluation of a short version of the quality of life questionnaire for dementia. *Int Psychoger* 2015; 27: 103-110.
15. Mahoney FI, Barthel DW. Functional evaluation: The barthel index. *Md State Med J* 1965; 14: 61-65.
16. Bonsignore M, Barkow K, Jessen F, et al. Validity of the five-item WHO Well-Being Index (WHO-5) in an elderly population. *Eur Arch Psychiatry Clin Neurosci* 2001; 251: 27-31.
17. Kunisada M. Research in relation with the care working hour of long term care works in welfare institute for elderly, their level of burden and level of achievement. *Japanese Society of Human Sciences of Health-Social Services* 2011; 17: 1-8.
18. Cabinet Secretariat. Report on the implementation status of the declaration of a state of emergency for the new coronavirus infectious disease. [cited 12 January 2023]. Available from: [https://corona.go.jp/news/pdf/houkoku\\_r031008.pdf](https://corona.go.jp/news/pdf/houkoku_r031008.pdf).
19. Coco K, Kangasniemi M, Rantanen T. Care personnel's attitudes and fears toward care robots in elderly care: A comparison of data from the care personnel in Finland and Japan. *J Nurs Scholarsh* 2018; 50: 634-644.