Training Independent Observers to Identify Behavioral Symptoms in Nursing Home Residents with Dementia using the Agitated Behavior Mapping Instrument

Esme Zediker[*,1], **Ellen McCreedy, Ph.D.**[*,1,2,3], **Natalie Davoodi, MPH**[3], **Vincent Mor, PhD**[1,2,3,5], **James L. Rudolph, MD**[2,3,4,5]

[1.] Center for Long-Term Care Quality & Innovation, Brown University School of Public Health, Providence RI

[2.] Center for Gerontology & Healthcare Research, Brown University School of Public Health, Providence RI

[3.] Department of Health Services, Policy, and Practice, Brown University School of Public Health, Providence RI

[4.] Division of Geriatrics and Palliative Medicine, Warren Alpert Medical School of Brown University, Providence RI

[5.] Center of Innovation in Long-Term Services and Supports, Providence VA Medical Center, Providence RI

Introduction:

Behavioral symptoms of dementia (BSD) among nursing home (NH) residents with Alzheimer's disease and related dementia (ADRD) are distressing to residents and staff.[1] Residents with BSD are more likely to receive antipsychotic medications which increase the risk of falls and death in this population.[2] One barrier to real-world evaluations of non-pharmacologic alternatives is a lack of pragmatic data on BSD. While NHs are required to regularly assess residents using the Minimum Data Set (MDS), the MDS behavioral items are often completed by staff who do not work closely with the resident; represent only a snapshot in time (one week per quarter); and may be influenced by staff normalization of behaviors. Previous work has demonstrated that the MDS behavioral measures significantly under-detect BSD compare to gold-standard staff interviews.[3] Measures of BSD occurring close in time to the intervention, and collected by independent observers, may reduce measurement biases.

Corresponding Author: Ellen McCreedy, PhD, Assistant Professor, Center for Gerontology and Healthcare Research, Department of Health Services, Policy, and Practice, Brown University, School of Public Health, 121 South Main Street, Suite 6, Providence, RI 02903, ellen_mccreedy@brown.edu, (401) 863-7345.
[*]Contributed equally to manuscript

Methods:

Music & MEmory: A Pragmatic TRial for NH Residents With ALzheimer's Disease, or METRIcAL, was a pragmatic trial of a personalized music intervention to help manage behaviors in persons living with ADRD from 54 NHs (27 treatment, 27 control).[4] Due to the potential concerns with the behavioral measures in the MDS, we used the Agitated Behavior Mapping Instrument (ABMI) to assess BSD. The ABMI is a structured observation tool in which a trained rater assesses the frequency of 14 behaviors over a three-minute period.[5] For each behavior, the rater indicates how many times the behavior occurred (up to ten times) over the observation period. The ABMI has previously been used in research to assess behaviors in NH residents with advanced ADRD, but the inter-rater reliability has only been reported for observers trained directly by the tool developer.[5, 6]

To facilitate BSD measurement for METRIcAL at geographically diverse NHs, we conducted a data collector training before each of three planned NH data collection visits. The full protocol has been published elsewhere.[4] Briefly, ten data collectors were recruited based on their proximity to participating NHs. Data collectors attended three, two-day trainings. The first two training sessions were held in-person. The last training was held virtually using conferencing software. Training specific to the ABMI involved the use of standardized medical actors performing three-minute, vignettes which highlighted specific BSD. The vignettes were rehearsed and coached by a geriatrician (JLR) and performed live for the in-person trainings and via video for the virtual training. An automated version of the ABMI was completed by the trainees and responses were recorded with Qualtrics.

To assess the consistency of behavioral scores for ten independent raters viewing the same vignettes at the same time, we calculated the intraclass correlation coefficient (ICC) using a two-way mixed-effects model with a fixed effect for observers. The ICC is an appropriate measure of inter-rater reliability when the outcome is continuous and raters are evaluating multiple vignettes.[7] We assessed the consistency in the ABMI total score as well as the ABMI behavioral and verbal behavior domains.[5] All analyses were conducted using Stata version 16.[8]

Results:

There was consistency in the total overall ABMI behavioral score among the observers (Table 1). The inter-rater reliability increased with each subsequent training (Baseline ICC: 0.75; Midpoint ICC: 0.84; Endpoint ICC: 0.98). While the domain ICCs were consistently high, the physical domain was slightly lower at Baseline (Baseline ICC: 0.66; Midpoint ICC: 0.99; Endpoint ICC: 0.98) than the verbal domain (Baseline ICC:0.79; Midpoint ICC:0.96; Endpoint ICC:0.98). Most errors occurred in counting. For example, if a resident was picking at themselves, observers correctly identify the behavior as "repetitive mannerisms," but had difficulty determining when the behavior started and stopped.

Discussion:

We sought to train non-clinical observers to consistently identify BSD in NH residents with dementia. Our training program used medical actors and demonstrated increased

agreement among observers over three training sessions. The consistently high ICC demonstrates that lay professionals can achieve agreement on behavioral observations. Our findings are particularly important given the limitations in the widely available measures of behaviors.[9, 10]

Acknowledgements:

Sponsor's role:

Funding:

References

1. Kandelman N, Mazars T and Levy A. Risk factors for burnout among caregivers working in nursing homes. J Clin Nurs 2018; 27: E147–E153. DOI: 10.1111/jocn.13891. [PubMed: 28543882]

2. Maust DT, Kim HM, Seyfried LS, et al. Antipsychotics, other psychotropics, and the risk of death in patients with dementia: number needed to harm. JAMA Psychiatry 2015; 72: 438–445. 2015/03/19. DOI: 10.1001/jamapsychiatry.2014.3018. [PubMed: 25786075]

3. Bharucha AJ, Vasilescu M, Dew MA, et al. Prevalence of behavioral symptoms: comparison of the minimum data set assessments with research instruments. J Am Med Dir Assoc 2008; 9: 244–250. 2008/05/07. DOI: 10.1016/j.jamda.2007.08.005. [PubMed: 18457799]

4. McCreedy EM, Gutman R, Baier R, et al. Measuring the effects of a personalized music intervention on agitated behaviors among nursing home residents with dementia: design features for cluster-randomized adaptive trial. Trials 2021; 22: 681. 2021/10/09. DOI: 10.1186/s13063-021-05620-y. [PubMed: 34620193]

5. Cohen-Mansfield J, Werner P and Marx MS. An observational study of agitation in agitated nursing home residents. Int Psychogeriatr 1989; 1: 153–165. 1989/01/01. DOI: 10.1017/s1041610289000165. [PubMed: 2491142]

6. Cohen-Mansfield J, Thein K, Marx MS, et al. Efficacy of nonpharmacologic interventions for agitation in advanced dementia: a randomized, placebo-controlled trial. The Journal of clinical psychiatry 2012; 73: 1255–1261. [PubMed: 23059151]

7. Laschinger HK. Intraclass correlations as estimates of interrater reliability in nursing research. Western journal of nursing research 1992; 14: 246–251. 1992/04/01. DOI: 10.1177/019394599201400213. [PubMed: 1561790]

8. Ho AT, Huynh KP, Jacho-Chávez DT, et al. Data science in Stata 16: Frames, lasso, and Python integration. Journal of Statistical Software 2021; 98: 1–9.

9. McCreedy EM, Yang X, Baier RR, et al. Measuring Effects of Nondrug Interventions on Behaviors: Music & Memory Pilot Study. J Am Geriatr Soc 2019; 67: 2134–2138. 2019/07/14. DOI: 10.1111/jgs.16069. [PubMed: 31301191]

10. McCreedy E, Ogarek JA, Thomas KS, et al. The Minimum Data Set Agitated and Reactive Behavior Scale: Measuring Behaviors in Nursing Home Residents With Dementia. J Am Med Dir Assoc 2019; 20: 1548–1552. 2019/11/05. DOI: 10.1016/j.jamda.2019.08.030. [PubMed: 31678075]

Table 1:

Inter-rater reliability achieved by lay observers watching three-minute vignettes performed by trained medical actors, as recorded using the Agitated Behavior Mapping Instrument

	Baseline Training (10 raters, 8 vignettes) ICC (95% CI)	Midpoint Training (10 raters, 2 vignettes) ICC (95% CI)	Endpoint Training (9 raters, 5 vignettes) ICC (95% CI)
ABMI Overall	0.75 (0.45, 0.95)	0.84 (−0.18, 1.00)	.98 (0.92, 1.00)
Physical Behaviors[a]	0.66 (0.16, 0.92)	0.99 [0.92, 1.00]	0.98 [0.93, 1.00]
Verbal Behaviors[b]	0.79 [0.48, 0.95]	0.96 [0.74, 1.00]	0.98 [0.93, 1.00]

Abbreviations: ABMI; agitated behavior mapping instrument; ICC; intraclass correlation; CI; confidence interval

[a]Physical behaviors include: hitting, kicking, pushing, scratching, tearing, cursing, grabbing, pacing, disrobing, exiting, handling things inappropriately, restlessness, repetitious mannerisms.

[b]Verbal behaviors include: attention-seeking, repetition, complaining, screaming, groaning, nonsense, unwarranted laughter, singing.