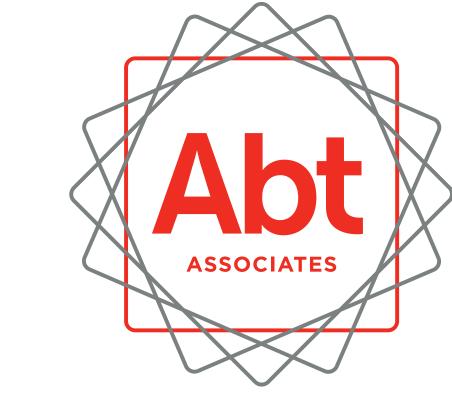
# Reliability and Validity Testing of Current and Potentially New OASIS Items







antana CONSULTING GROUP

Eugene Nuccio, PhD¹; Jennifer Riggs, PhD RN²; Zabrina Gonzaga, MSN RN³; Alan Levitt, MD⁴; Tara McMullen, PhD⁴; Joan Proctor, MS HCA⁴; David Hittle, PhD¹; Jacqueline Gillis²

<sup>1</sup>University of Colorado Anschutz Medical Campus; <sup>2</sup>Abt Associates; <sup>3</sup>Lantana Consulting Group; <sup>4</sup>Centers for Medicare & Medicaid Services, Division of Chronic and Post-Acute Care (DC-PAC)

### Background

- Home health agencies (HHAs) are required to report data, including Outcome and Assessment Information Set (OASIS) data, to CMS in support of the Home Health Quality Reporting Program that documents the HHA's risk adjusted performance in the care of their patients.
- Selected OASIS-based quality performance measures are publicly reported on the Medicare.gov website Home Health (HH) Compare and include the HH Quality of Patient Care star ratings and the HH Value-Based Purchasing Model.
- One of the overarching goals of the OASIS Field Test was to assess the sensitivity, reliability and validity of existing functional and selected functional standardized patient assessment data elements (SPADE).
- Empirical evidence of item and measure sensitivity, reliability and validity is required as part of new and continuing National Quality Forum endorsement.

### Research Objective

The primary purpose of the OASIS Field Test was to assess the sensitivity, reliability, validity, and feasibility of selected existing OASIS and SPADE items.

### Learning Objectives

- Compare sensitivity and reliability differences between existing OASIS and SPADE items.
- Assess validity and feasibility of selected existing OASIS and SPADE items.

### Population, Study Design, Methodology

- HH patients (n=213) from 12 Medicare-certified home health agencies in four states participated in the field test.
- Field test team used quantitative and qualitative data collection and analysis techniques to evaluate the reliability and validity of existing and potentially new OASIS items.
- Pairs of trained assessors (registered nurses and physical therapists) collected field test data within 24 hours of each other either at Start of Care (SOC) or discharge.
- Interrater reliability was calculated using linear weighted kappa to account for the magnitude of difference in paired assessor ratings.
- Correlations between raters were also computed. Focus groups, clinician surveys and medical record reviews were used to assess validity and feasibility.

### **Correlations and Comparison of** Current OASIS vs. SPADE Items

- Correlations between raters at SOC on existing OASIS and SPADE items (physiological (not shown) and functional (shown)) were statistically significant (p<0.001).
- Correlations between raters were also extremely high at discharge.
- Given that a high percentage of patients were assessed as independent at both SOC and discharge using the SPADE items, the high correlations for this group of items may be misleading.

#### Comparison of Independence between **Current OASIS and SPADE Functional Items**

- Current OASIS Functional Items
- The median independence at SOC was in the low 20% range, with only two items (Eating & Management of Oral Medications) exceeding 40% independence at SOC.
- Only 3 items (Toilet Transferring; Toileting Hygiene; and Eating) of the 10 current OASIS functional items exceeded 80% at discharge.
- SPADE Functional Items
- The median independence at SOC was in the low 40% range, and three items (Eating, Oral Hygiene, Roll Left/Right) reported SOC independence rates between 65-75%.
- All 11 of the SPADE functional items exceeded 80% at discharge.

## Results Overview

#### Comparison of Kappa Values (Current OASIS vs. SPADE Functional Items)

- Current OASIS functional items:
- Wt'd Kappa values for existing OASIS items typically rated "moderate" at 0.49 for SOC and were slightly higher ("substantial") at discharge (mean = 0.575; median = 0.59).
- At SOC two items were rated as "fair or worse" while at discharge only one item was rated "fair or worse".
- SPADE functional items:
- Wt'd Kappa values for potentially new OASIS "GG" items were typically rated "fair" (mean = 0.388; median = 0.44) at SOC, while there the values were higher at discharge (mean = 0.555; median = 0.605).
- At SOC five items were rated "fair or worse" while at discharge four items were rated "fair or worse".

GG0170J Walk 50 ft.

#### Clinician Evaluation

- Clinician survey and focus group data supported validity and feasibility of some potentially new items tested, including those for pain and cognition.
- Questions raised included
- highlighting opportunities for focused training, and
- the need for specific guidance in the form of a Manual on scoring of potential new items that have a reverse scale from current OASIS items.
- One additional validity finding based on analytic results was that potentially new OASIS items for communication and continence produced comparable results for parallel current OASIS items.
- Differences in the field test performance between existing OASIS and SPADE items may indicate that HHA-level training is needed prior to beginning the data collection with the new OASIS-D instrument.

items are introduced.

Conclusions

Findings suggest all items tested showed reasonable,

The SPADE functional items showed less sensitivity

independent at SOC) than current OASIS functional

independence across all items at discharge when

compared with a much wider range of independence

reported at discharge for current OASIS functional

OASIS) and unfamiliar (SPADE) items at both SOC

and discharge suggest the need for agency-level

training and detailed guidance as new assessment

Limitations

because their patients were recruited based on ability

The participating HHAs are a convenience sample

(i.e., higher percentages of patients rated

items did, as well as nearly uniform ratings of

Differences in values between familiar (existing

albeit not equal, levels of reliability.

 Qualitative data support validity and feasibility of most items tested, although the need for focused training and guidance were identified by clinicians.

#### **Correlations for** Current OASIS and SPADE functional items

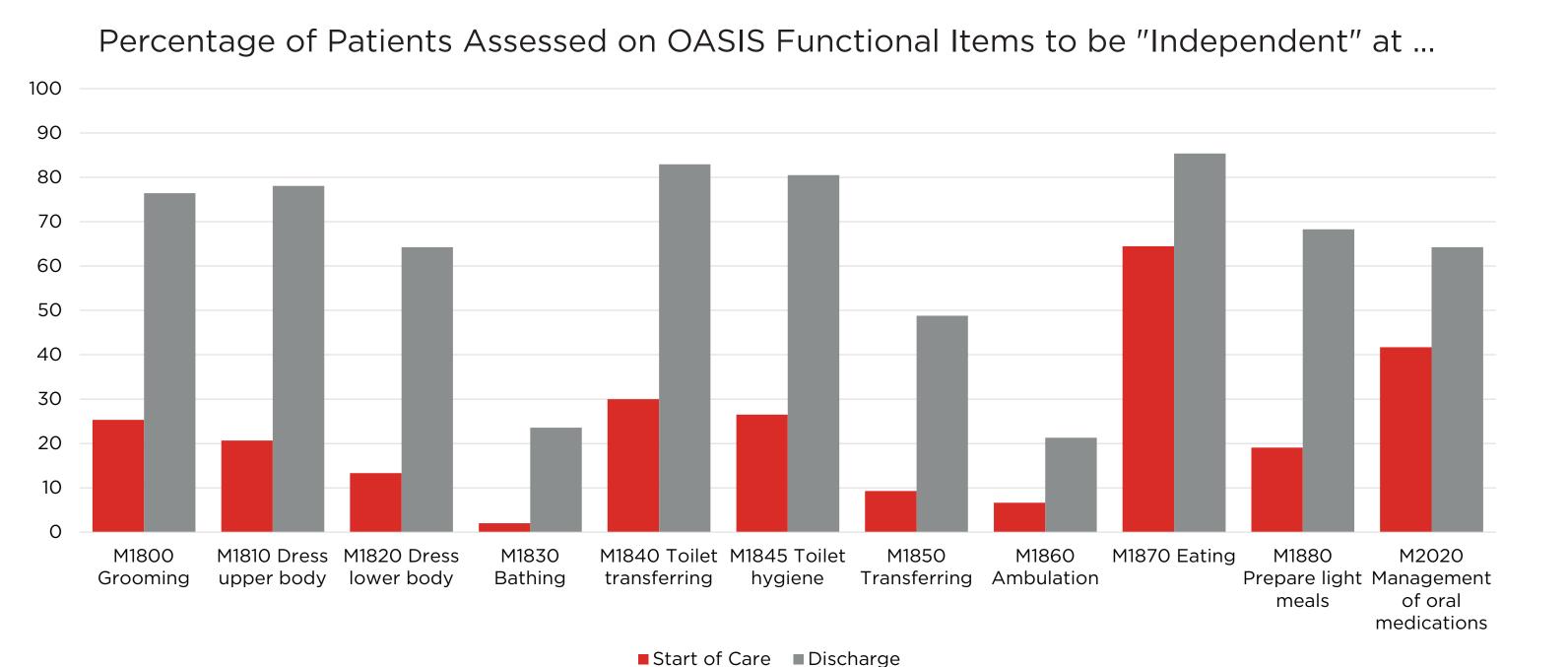
OASIS Item	n	n <sub>R</sub>	Mean	Mean <sub>R</sub>	Corr.	p-value
M1800 Grooming	104	104	1.00	1.01	0.34	0.0004
M1810 Dress upper body	104	105	1.15	1.14	0.51	<.0001
M1820 Dress lower body	104	105	1.63	1.56	0.58	<.0001
M1830 Bathing	104	105	2.86	2.81	0.52	<.0001
M1840 Toilet transferring	104	105	0.86	0.90	0.49	<.0001
M1845 Toilet hygiene	104	105	1.03	1.14	0.52	<.0001
M1850 Transferring	104	105	1.31	1.31	0.52	<.0001
M1860 Ambulation	105	104	2.54	2.60	0.48	<.0001
M1870 Eating	105	105	0.44	0.40	0.20	<.0001
M1880 Prepare light meals	105	105	1.21	1.13	0.47	<.0001
M2020 Management of oral meds.	105	105	1.17	1.17	1.00	<.0001

SPADE Item	n	n <sub>R</sub>	Mean	Mean <sub>R</sub>	Corr.	p-value
GG0130 Eating	103	101	5.69	5.66	0.21	0.0384
GG0130 Oral hygiene	101	99	5.44	5.38	0.53	<.0001
GG0130 Toileting hygiene	103	103	4.74	4.78	0.46	<.0001
GG0130 Wash upper body	99	102	4.71	4.85	0.38	0.0001
GG0170A Roll left and right	98	96	5.26	5.26	0.51	<.0001
GG0170B Sit to lying	101	99	4.84	4.98	0.53	<.0001
GG0170C Lying to sitting	101	98	4.70	4.77	0.59	<.0001
GG0170D Sit to stand	101	105	4.43	4.58	0.49	<.0001
GG0170E Chair transfer	103	103	4.43	4.21	0.57	<.0001
GG0170F Toilet transfer	103	100	4.45	4.58	0.66	<.0001
GG0170I Walk 10 ft.	99	96	4.67	4.46	0.44	<.0001
GG0170J Walk 50 ft.	85	80	4.19	4.39	0.52	<.0001

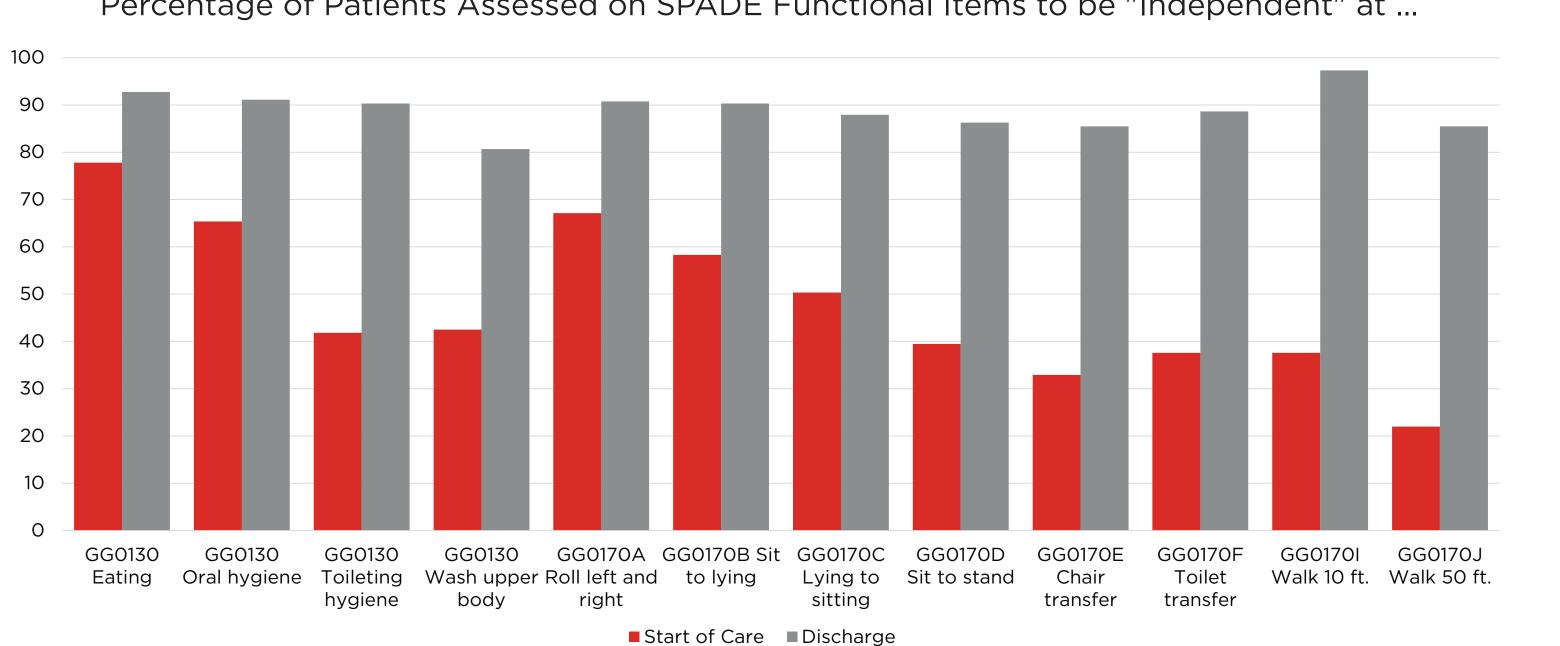
MeanR = Mean at Discharge

nR = n at Discharge

#### **SOC & Discharge Percent of Patients** Assessed as Independent on OASIS Items



SOC & Discharge Percent of Patients Assessed as Independent on SPADE Items Percentage of Patients Assessed on SPADE Functional Items to be "Independent" at ...



### Wt'd Kappa Values: **Current OASIS Functional Items at SOC**

Current OASIS Item (#, Title at SOC)	N1	Weighted Kappa	Wt'd Kappa Rating	% Observed Agreement
M1800 Grooming	104	0.37	Fair	60.8%
M1810 Dress upper body	104	0.51	Moderate	66.0%
M1820 Dress lower body	104	0.58	Moderate	65.0%
M1830 Bathing	104	0.51	Moderate	58.3%
M1840 Toilet transferring	104	0.49	Moderate	68.9%
M1845 Toilet hygiene	104	0.51	Moderate	53.4%
M1850 Transferring	104	0.42	Moderate	66.0%
M1860 Ambulation	105	0.43	Moderate	67.0%
M1870 Eating	105	0.22	Fair	62.5%
M1880 Prepare light meals	105	0.41	Moderate	60.6%
M2020 Management of oral meds.	105	1.00	Perfect	100.0%

7.4%

**Current SPADE Item** % Observed (#, Title at SOC) Rating2 **Agreement** GG0130 Eating 63.3% GG0130 Oral hygiene 62.8% 21.0% GG0130 Toileting hygiene 22.1% GG0130 Wash upper body GG0170A Roll left and right 51.1% 0.23 37.9% GG0170B Sit to lying GG0170C Lying to sitting 14.9% GG0170D Sit to stand GG0170E Chair transfer 11.0% 19.6% GG0170F Toilet transfer GG0170I Walk 10 ft. 99 47.3%

85

0.29

Fair

Wt'd Kappa Values:

SPADE Functional Items at SOC

Slight = 0.10 to 0.20 Fair = 0.21 to 0.40 Moderate = 0.41 to 0.60 Substantial = 0.61 to 0.80 Near perfect= 0.81 to 0.99 Perfect = 1.00

### Findings are based on a small sample size and may

not be generalizable to the home health patient population nationally. The training provided to the assessors on using the

to participate and study team's location.

#### SPADE items that have a standardized and reversed rating scale when compare with the existing OASIS items may have been inadequate and may be related to differences in item sensitivity and reliability.

### Implications for Policy and Practice

- Results of these analyses can inform overall strategic discussions about item retention, replacement, or removal as CMS transitions to standardized data elements across post-acute care settings.
- Findings inform ongoing efforts to develop effective and efficient training and guidance for clinicians tasked with collecting data on OASIS items.

#### **Funding Acknowledgement**

Funding was received from the Centers for Medicare & Medicaid Services through Contract are solely those of the authors and do not reflect the official positions of the institutions of organizations with which they are affiliated or the views of the project sponsors.



