RAND HEALTH

Note: If you prefer to view this as a document for use with a text reader, please see alternate document.

Special Open Door Forum

January 24, 2008

Main Advances in MDS 3.0

- **Gives Resident Voice**
- **Increases clinical relevance**
- **Increases accuracy (validity & reliability)**
- **Increases clarity**
- **Reduces time to complete by 45%**

Why Resident Voice?

- **CMS's goal is to increase resident-centered care**
  - **Respect for individual voice**
  - **Fundamental to high quality & culture change**
  - **Residents and families want care to be individualized and accurate**
- **Improves accuracy, feasibility, efficiency**
  - **General, unfocused questions do not elicit meaningful reports**
  - **Detailed daily observations of all behaviors for all residents is time consuming and not feasible**

RAND HEALTH

How did we identify and test these advances?

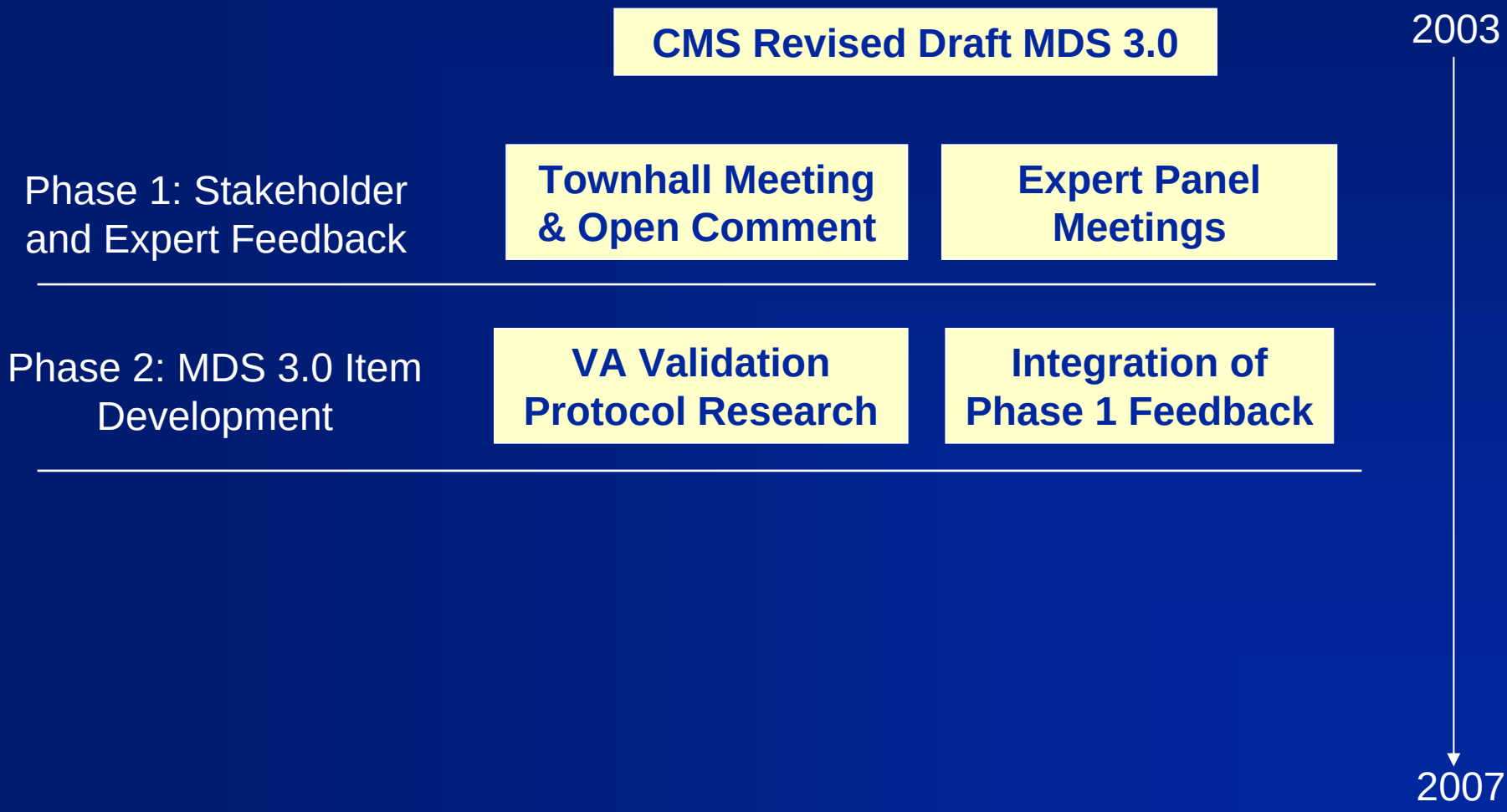
The Evaluation Team Had 6 Sets of Players

- **Lead research and administrative team**
  - RAND: Debra Saliba, MD, MPH
  - Harvard: Joan Buchanan, PhD
  - Administrative Lead: Malia Jones
- **National VA Nursing Home Research Collaborative**
  - Los Angeles, CA
  - Philadelphia, PA
  - Atlanta, GA
  - Bedford, MA
- **Lead Quality Improvement Organization:**

  Colorado Foundation for Medical Care
- **Instructions, Guides and Form Design:**

  Carelink, RRS Consulting, Kleimann Communications Group
- **Centers for Medicare & Medicaid Services**
- **Workgroups, consultants, content experts**

MDS 3.0 Development Proceeded in 4 Phases



Phase 2 Improved Key MDS Sections and Revised MDS Items

The VA Pilot Developed and Tested MDS Items in 8 Areas

- **Mood**
- **Behavior disorders**
- **Mental status**
- **Delirium**
- **Pain**
- **Falls**
- **Quality of life**
- **Diagnostic coding**

Funded by VA HSR&D

Findings of VA Research

Depression

- Self-report is feasible & efficient
- Yields more valid estimates

Pain

- Self-report is feasible & efficient
- Yields more valid estimates than observation
- Ascertaining impact on function is feasible and provides useful information

Customary Routine and Activities

- As recommended by TEP and Validation panel, asking importance is feasible

Funded by VA HSR&D

Findings of VA Research

Diagnoses

- Algorithms to define active diagnosis improve identification compared to administrative data

Delirium

- Revised protocol and instruction improved agreement

Cognition

- Structured interview is feasible and welcomed by staff

Falls

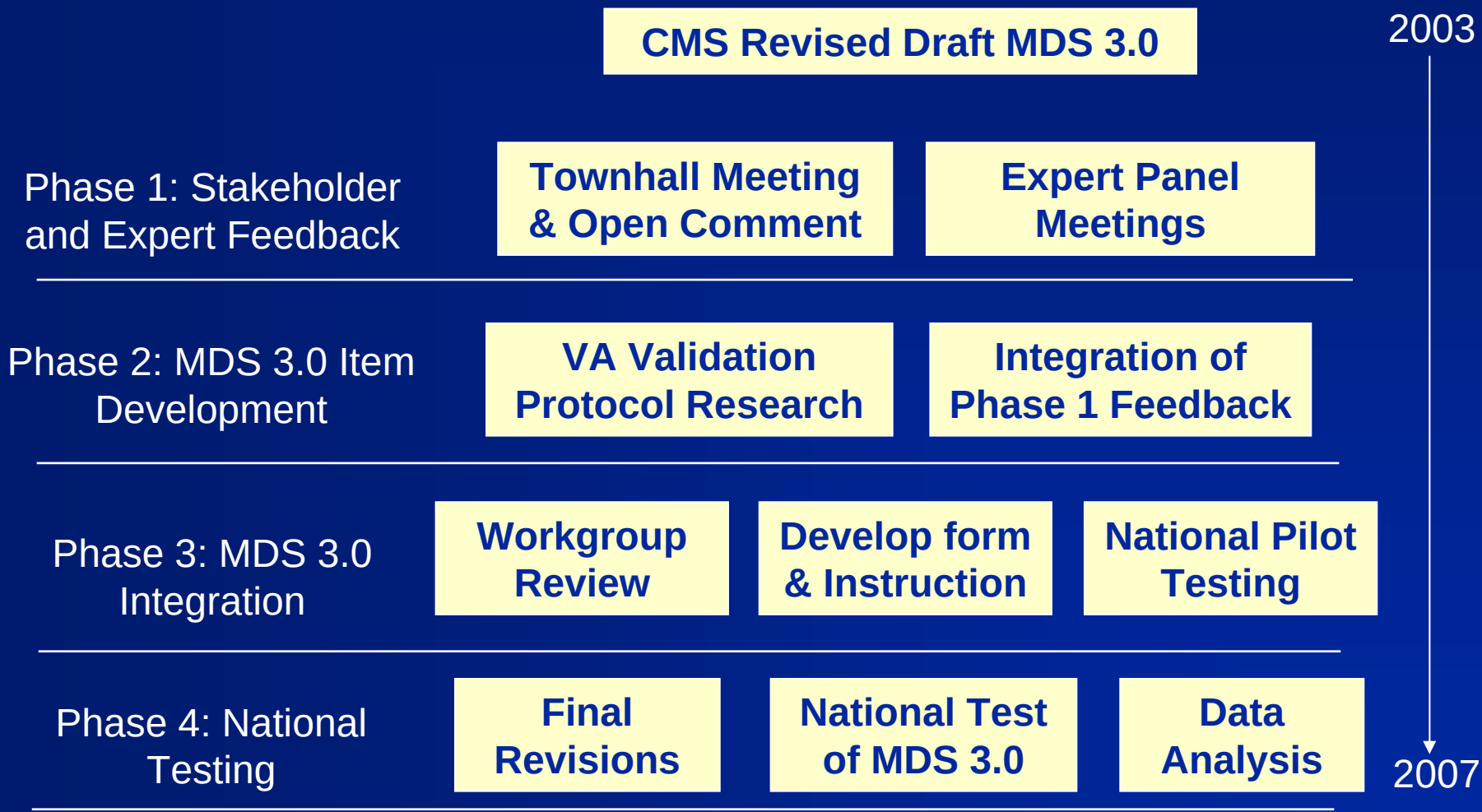
- Simplified response options can be used by NH staff to classify falls

Behavior

- Items can consider impact on resident

Funded by VA HSR&D

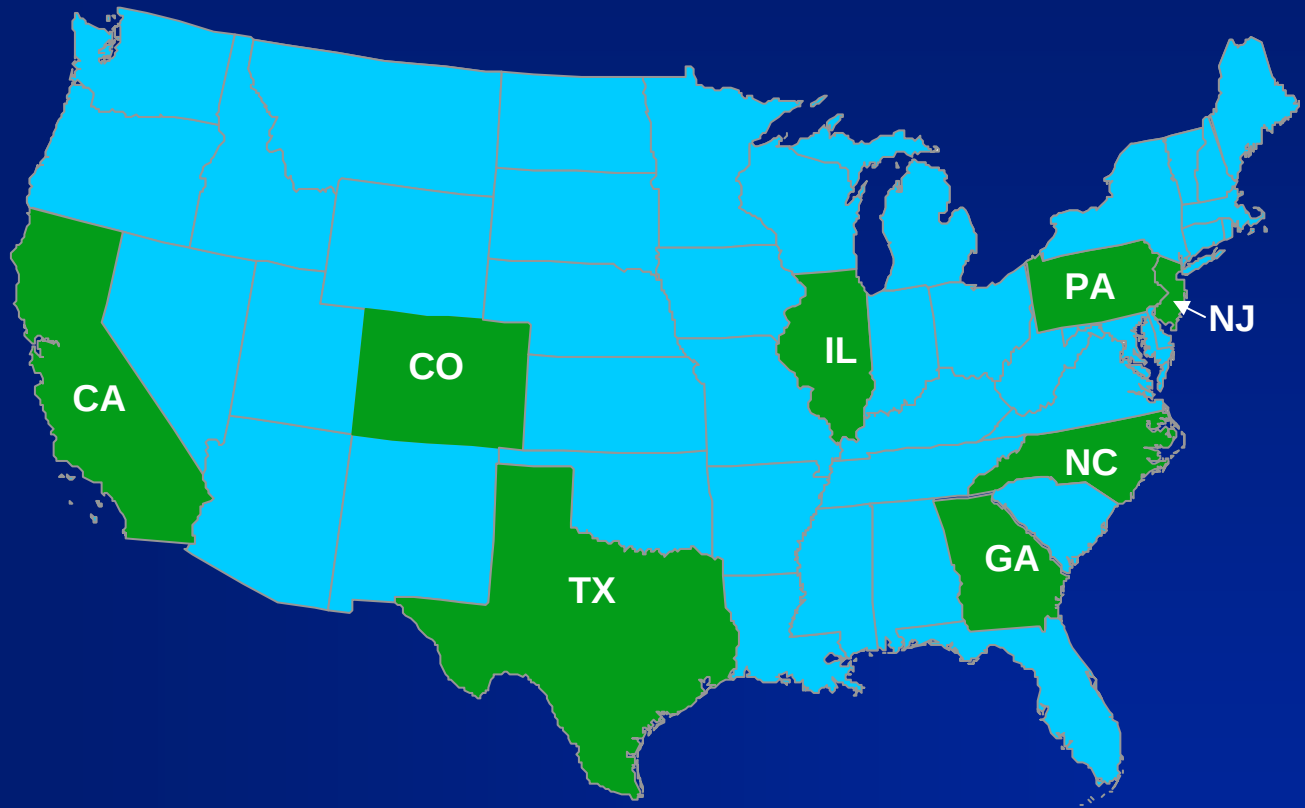
MDS 3.0 Development Proceeded in 4 Phases



Revised Form Design

- **Form structured to improve usability**
  - **Important definitions put on form**
  - **Larger font**
  - **Logical breaks, fewer items to a page**
- **Items that were confusing or not needed for programming deleted**

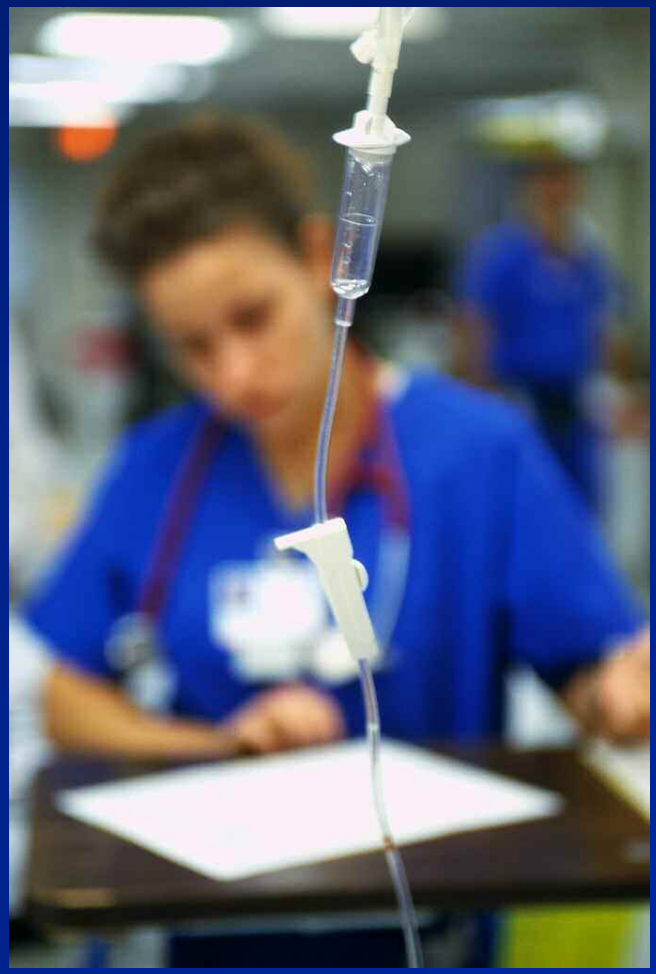
MDS 3.0 Was Tested 71 NHs in 8 States



3800 residents participated in different parts of the evaluation

Different Types of Data Collectors Tested MDS 3.0

- **Each state had 2 gold standard nurse data collectors**
- **Each nursing home had 1 facility nurse data collector**



The National Test Measured Reliability and Validity of MDS 3.0

Reliability of MDS 3.0

- Inter-rater reliability measures the extent to which two data collectors achieve the same results when assessing the same event
  1. Gold-standard to Gold Standard
  2. Gold-standard to Facility Nurse

Validity of MDS 3.0

- Validity assesses the degree to which items measure the intended concept
  - content
  - criterion
  - construct

Additional Evaluation Elements

- Time to complete
  - Recorded all start and stop times for both MDS 2.0 and MDS 3.0
- Two Anonymous Nurse Surveys
  - Mailed to all nurses who participated
  - MDS 2.0 survey first
  - MDS 3.0 survey completed at end of study
  - Provided feedback on
    - Clinical usefulness of measures
    - Clarity and ease of completion
    - Satisfaction with assessment instrument
- MDS 2.0 collected to allow cross walk between instruments and into payment cells

RAND HEALTH

Review of 5 Sections with Major Revisions

1. MDS 3.0 Cognitive Assessment

Brief Interview for Mental Status (BIMS)

- New structured test replaces staff assessment for residents who can be understood

Staff Assessment for Mental Status

- Only completed for residents who cannot complete interview

Validated Confusion Assessment Method (CAM)

- Replaces old delirium items

Rationale for Cognitive Changes

Old cognitive item:

- **Providers express discomfort with observation-based scoring**
  - "long term memory OK" and "short term memory OK" items are not recognized by most providers
  - Only 29% thought MDS 2.0 easy to complete accurately
- **Instructs to use a formal assessment, but does not provide assessment or cross walk from standard assessment to 2.0**
- **CPS and COGs scales are not readily completed by NH staff**

New cognitive item:

- **Directly tests domains common to most cognitive tests in other settings –registration, temporal orientation, recall**
  - Partial credit for close answers & response to prompts makes more relevant for population
- **Supports validated delirium assessment protocols**

Rationale for Delirium Changes

Delirium is a serious condition associated with increased mortality, morbidity, costs and institutionalization

Old delirium items:

- **Reliability in some studies worse than chance**
- **Independent evaluations show significant under-detection with unstructured observation**

New delirium items = Confusion Assessment Method (CAM)

- **CAM is cited as appropriate tool by Royal College of Physicians, NCQA, other guidelines**
- **Improved sensitivity & specificity for detecting delirium**

BIMS Feedback Survey Results

- **80% thought BIMS improved ability to calculate score and trigger RAPs**
- **78% preferred BIMS interview to old assessment items**
- **88% reported that BIMS provided new insights into resident's cognitive abilities**

Results: Cognitive Item Performance

- **BIMS showed excellent reliability (kappa for score = .95)**
- **Completion rates were high**
  - **85% of residents were able to complete**
    - **Scores ranged from 0-15**

BIMS had excellent performance as a test to detect impairment

- **BIMS was more highly correlated with gold-standard measure**
  - **MDS. 3.0 BIMS = 0.91 (< .0001)**
  - **MDS 2.0 CPS = - 0.74 (<.0001)**
- **BIMs had a higher area under the receiver operating characteristics curve (AUC) for detecting impairment**
  - **BIMS AUC = .930**
  - **CPS AUC = .824**

(AUC: 1 = a perfect test .5= worthless)

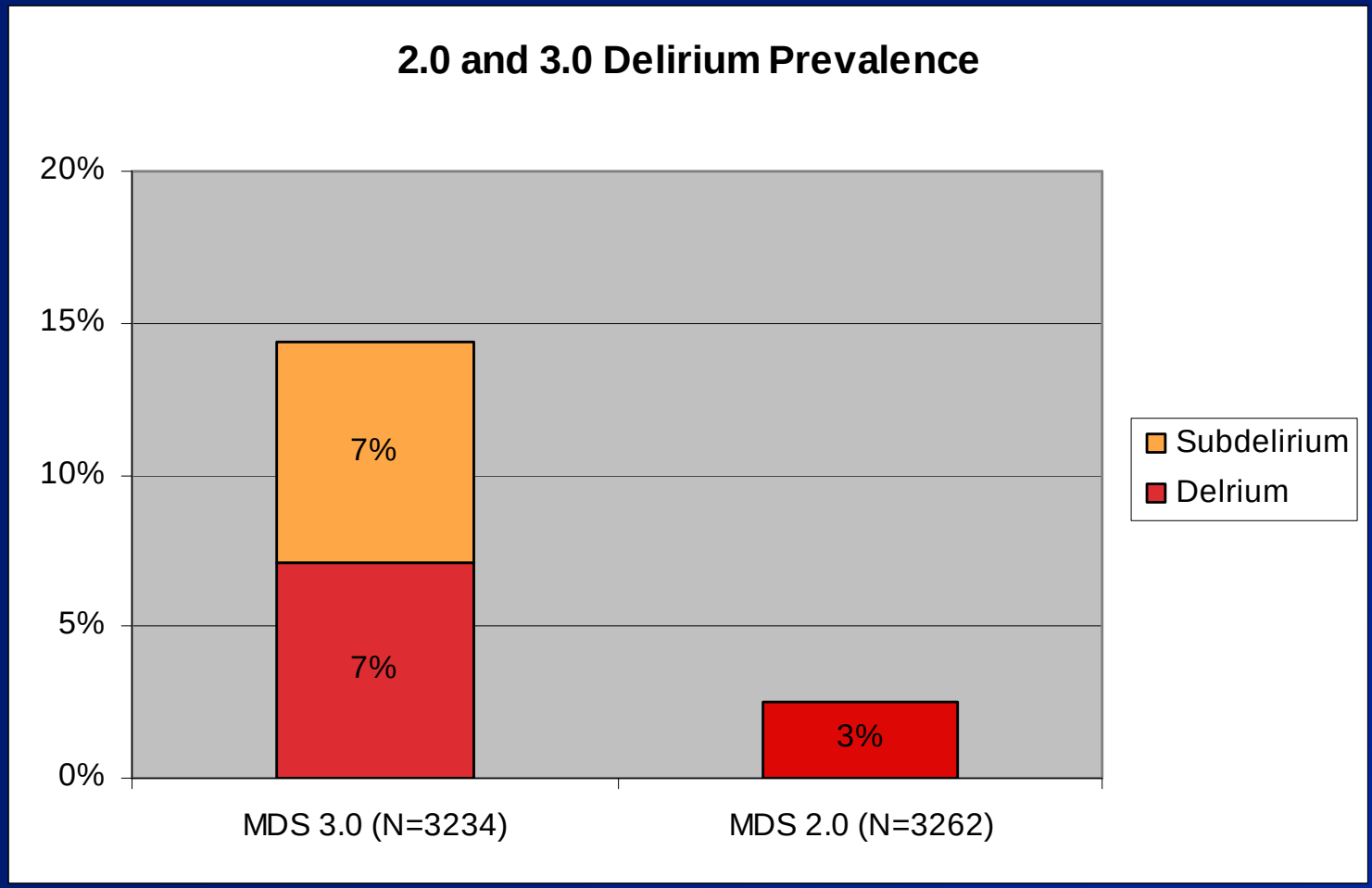
Delirium Feedback Survey Results

- **85% found definitions on form clear**
- **71% felt that CAM helped them do a better job of screening for delirium** (7% disagreed)
- **64% reported that BIMS led them to observe new delirium behaviors that differed from those in medical record**

Delirium showed very good reliability

- **Item reliabilities ranged from kappa = .75 to .89**

Delirium prevalence more consistent with expected rates



2. MDS 3.0 Mood Assessment

PHQ-9

New resident interview replaces staff observations for residents who can report mood symptoms

Staff Assessment of PHQ-9-OV

New observational items replace old staff assessment and only completed for residents who cannot self-report

- Includes irritability item

Rationale for Replacing Mood Items

Old mood item:

- **Repeatedly shown to have poor correspondence with independent mood assessment**
  - **Does not comport with accepted standard of self-report**
  - **Requires time consuming systematic observations of ALL residents across all shifts. Difficult to achieve.**
    - Only 22% reported that 2.0 section was easy to complete accurately
- **Questionable utility for gauging response to treatment, since appropriate approach is targeting DSM-IV signs and symptoms**

Rationale for Replacing Mood Items

New mood item (PHQ-9)

- Based on DSM-IV criteria
- Validity well established in other settings
- Increasing use and recognition by clinicians
- Allows threshold definition AND rapid sum of a severity score that can track change over time
- Has been used in outpatient elders, hospital, rehabilitation (post stroke) and home health populations in addition to younger adult populations

Mood Feedback Results

- 87% nurses rated the mood section as improved over 2.0 section
- 88% felt PHQ-9 interview was better than 2.0 observation for capturing resident mood
- 84% felt the items could inform care plans
- 86% reported that items provided new insights into mood

Feedback
Staff Mood Assessment

- **90% felt that detection and communication about mood would improve if staff learned to watch for these signs and symptoms**
- **72% found PHQ-9-OV assessment easier than MDS 2.0**

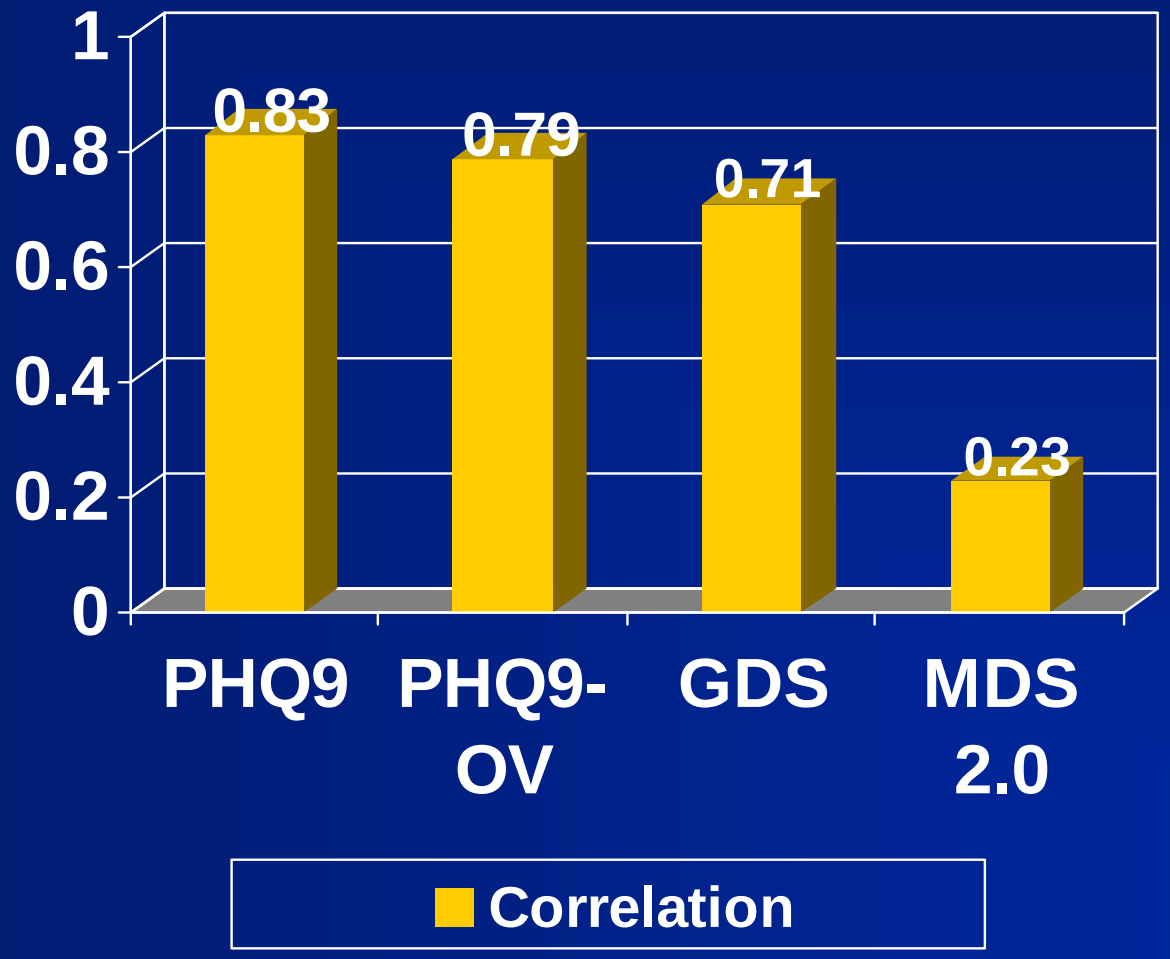
PHQ-9 showed excellent reliability

- **Resident Mood Interview**
  - **Kappa: 0.94**
- **Staff Mood Observations**
  - **Kappa: 0.93**

Ability of Residents to Complete PHQ-9

- **82% of non-comatose residents were able to complete interview**

PHQ-9 interview had best agreement with Gold Standard



3. MDS 3.0 Behavior Items

- Hallucinations and psychosis
  - moved from checklist in section J & definitions put on form
- Behaviors
  - Revised language clearer, linked to operational definitions
  - Revised symptom groupings to match constructs
  - Replaced “alterability” with specific impact questions
  - Replaced “resisting care” with “reject care” and refocused on resident’s goals of care
- Wandering rated separately from the 3 behavioral symptoms groups, and impact replaces alterability

Rationale For Changes

- Old behavior item groupings were not consistent with recognized factors
  - Only 41% of nurses rated MDS 2.0 items as easy to complete accurately
- Old behavior item labels were viewed as pejorative by consumers, did not convey potential expression of unmet need
  - New labels agreed to by providers & consumers
- Staff varied widely in definition of “alterability”
  - Alterability does not distinguish ongoing behaviors that require intervention
- New specific impact items give insight into severity and potential need for treatment/intervention

Behavior Feedback Survey Results

- **90% rated behavioral symptoms as easy to complete accurately**
- **91% nurses preferred the 3.0 behavior item section** (1% disagreed)
- **90-94% rated new behavioral symptoms items as clear**
- **88% rated impact items as providing important severity information**

Behavior section reliability was excellent

- **Psychosis, kappa = 0.96**
- **Overall kappa for all other behavioral items = 0.94**

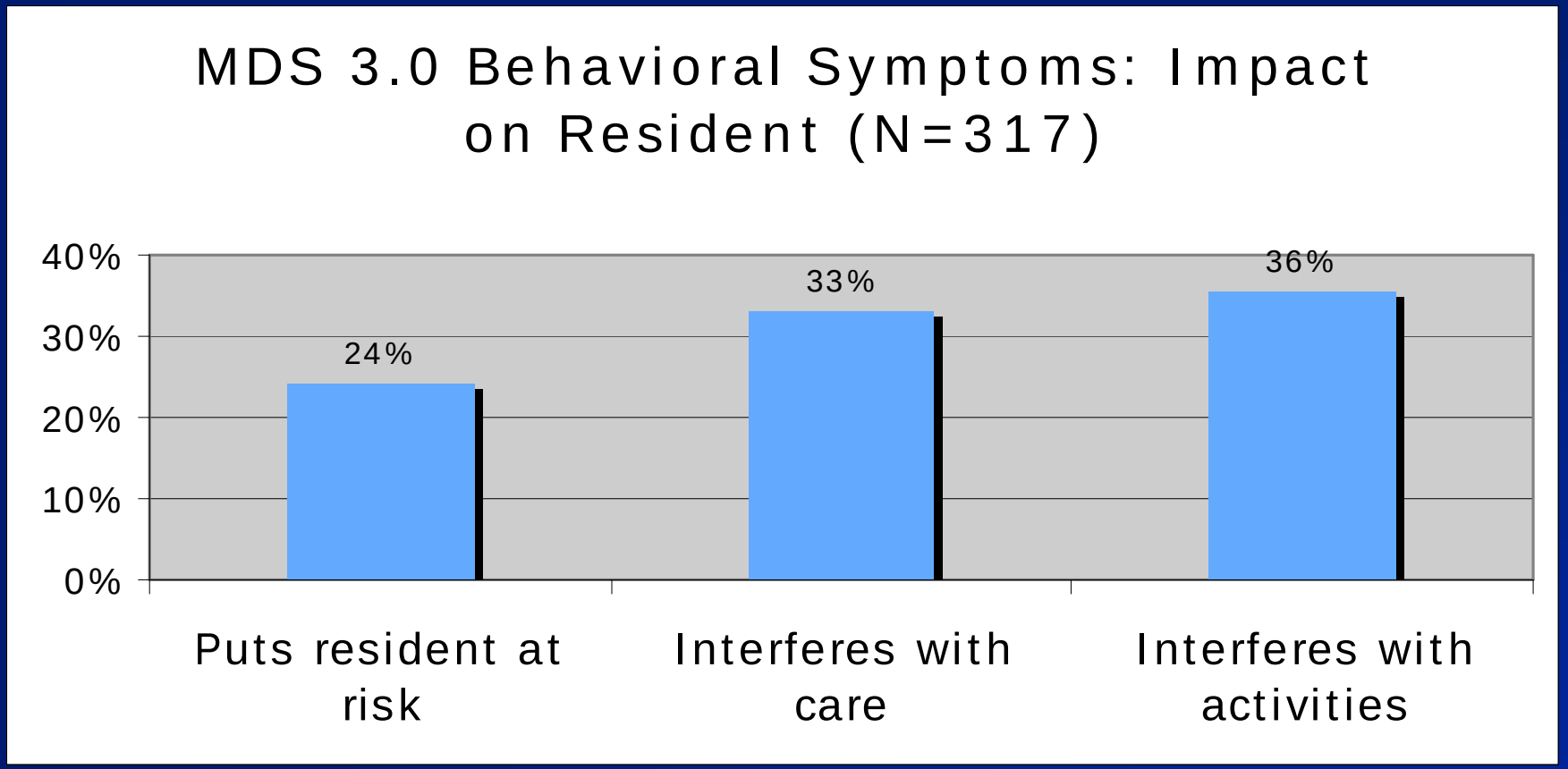
MDS 3.0 had Significantly Stronger Agreement with Gold Standard

CMAI Factor	MDS 3.0 Kappa (95% CI)	MDS 2.0 Kappa (95% CI)
Physical toward others	.86 (.74, .97)	.23 (.03, .43)
Verbal toward others	.73 (.61, .84)	.31 (.16, .45)
Other	.53 (.42, .66)	.22 (.12, .31)

MDS 3.0 Psychoses item also had Stronger Agreement with Criterion

NPI, Presence	MDS 3.0 Kappa (95% CI)	MDS 2.0 Kappa (95% CI)
Hallucinations	.92 (.81, 1.00)	.23 (.03, .43)
Delusions	.88 (.79, .98)	.31 (.16, .45)

Type of Impact on Resident Varies



4. MDS 3.0 Customary Routine & Activities

- **Preferred Routine**
  - New interview replaces 20 Customary Routine staff assessment items for residents who can be interviewed
  - Current importance rating replaces check all that apply in past year
- New interview for activities preferences replaces 12 staff assessment items for residents who can be interviewed
- **Want to talk to someone about returning to community**
- **Staff Assessment of Activity and Daily Preferences**
  - Only completed for residents who cannot complete interview
  - Major changes to several items; instructed to observe resident response during exposure to activity

Rationale for changes

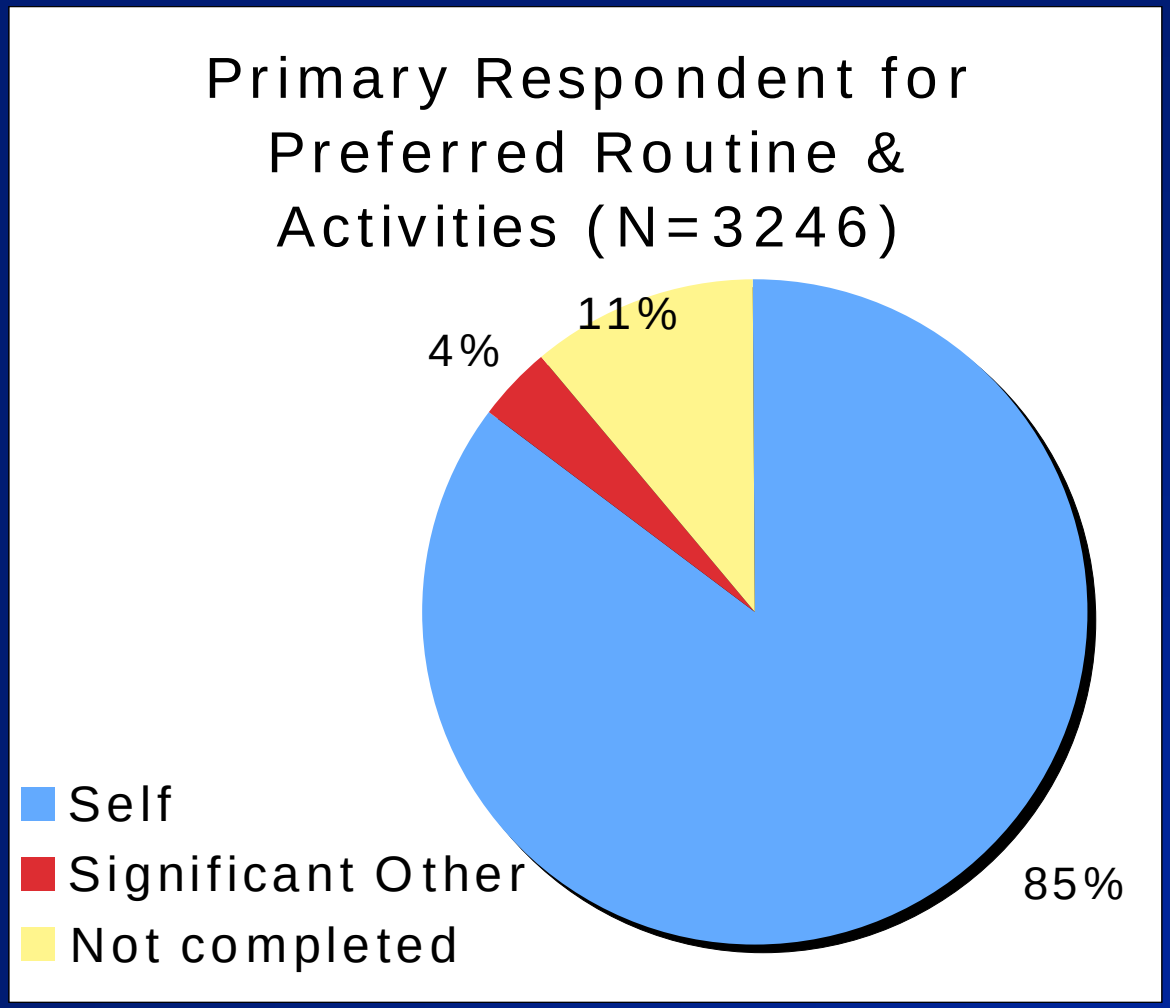
Old items

- Not perceived as helping with care planning
  - Prior practice could be related to ability, illness, access, not to preference
  - Only 30% rated 2.0 as helping care planning
- TEP and Validation Panels both recommended replace with importance scales

New Items (Preference Assessment Tool or PAT)

- Grounded in residential care quality
- Map to U Minnesota QoL domains
- Focuses on resident as central to determining activities

Residents were able to complete



Customary Routine Feedback Results

- 81% rated the interview items as more useful for care planning
- 80% found that the interview changed their impression of resident's wants
- More likely to report that post-acute residents appreciated being asked
- Only 1% felt that some residents who responded didn't really understand the items

Activity Feedback Results

- **77% rated as more useful for care planning** (1% disagreed)
- **83% found that the interview changed their impression of resident's wants**
- **Equally likely to report that post-acute residents appreciated being asked**
- **90% responded that residents answering questions understood** (0% disagreed)

Customary Routine & Activities Agreement was Excellent

- **Preferred Routine**
  kappa = 0.97
- **Activities 0.96**
  kappa = 0.96
- **Staff assessment of preferences:**
  kappa = 0.93

Overall scores were similar across cognitive groups for daily routine

Cognitive group	Range	Mean (sd)
Intact (n=1384)	**0 - 4**	**2.44 (1.08)**
Impaired (n = 734)	**0 - 4**	**2.60 (1.05)**
Severely impaired (n=827)	**0 - 4**	**2.46 (1.09)**

Same pattern was seen with activity items

5. MDS 3.0 Pain Assessment Items

- Treatment items added
- Resident interview replaces staff observations for residents who can report pain symptoms
- Section expanded to capture effect on function
- Staff assessment of pain changed to an observational checklist of pain behaviors and only completed for residents who cannot self-report

Rationale for Replacing Pain Items

Old pain item

- **Repeatedly shown to have poor correspondence with independent pain assessments**
  - **Does not comport with accepted standard of self report**
  - **Requires time consuming systematic observations of all residents across all shifts**
  - **Detection bias penalizes more vigilant facilities**
- **Providers and consumers frustrated that section addresses limited characteristics, insufficient to capture pain experience**
  - **3 point severity response insufficient and not match commonly used pain scales. Want severity response between "moderate" & "horrible or excruciating"**

Rationale for New Pain Items

- CMS and providers requested items to capture therapy
- Self-report is the gold standard for pain assessment
  - Pilot test showed ability to recall over 5 days
- With pain being reported as "5th vital sign" providers have increasingly used 0-10 scales in NHs & other settings
  - 0-10 scale would allow comparison across settings

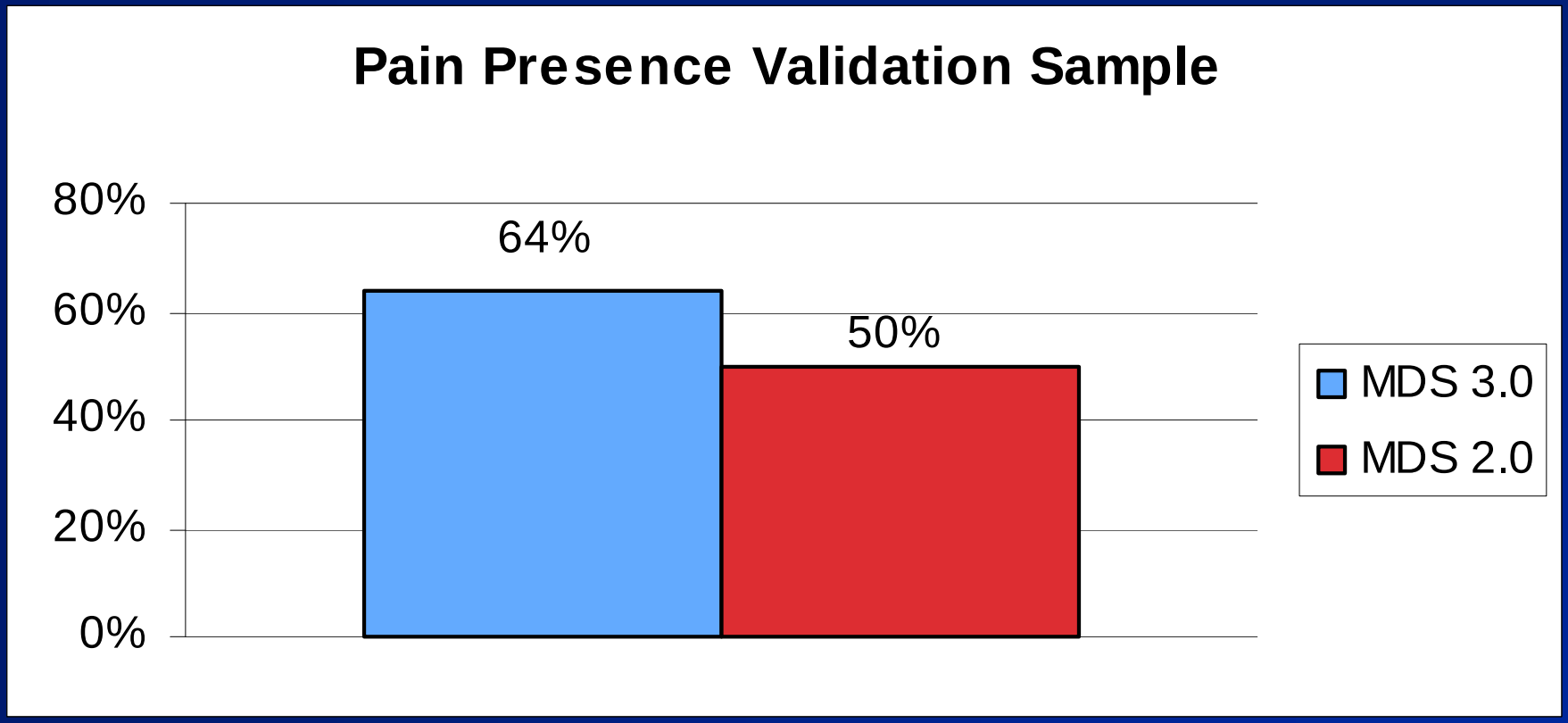
Pain Feedback was Positive

- 91-97% of nurses rated pain management item definitions clear
- 88% rated MDS 3.0 pain items as improved over MDS 2.0
- 94% reported that new pain items could inform care plans
  - Even during testing, pain interview provided new insights into resident's pain (85%)
- 90% felt that all residents who responded, understood (3% disagreed)
- 85% felt the observable behaviors would improve reporting of possible pain

Results Pain Item Performance

- **Pain Items showed excellent reliability**
  - **Pain treatment & interview (J1-J7)**
    - **kappa = 0.92**
  - **Staff assessment of pain (J9)**
    - **kappa = 0.97**
- **Completion rates were high**
  - **85% of non-comatose residents were able to complete the pain interview**

Pain Presence



Other Measures of Agreement

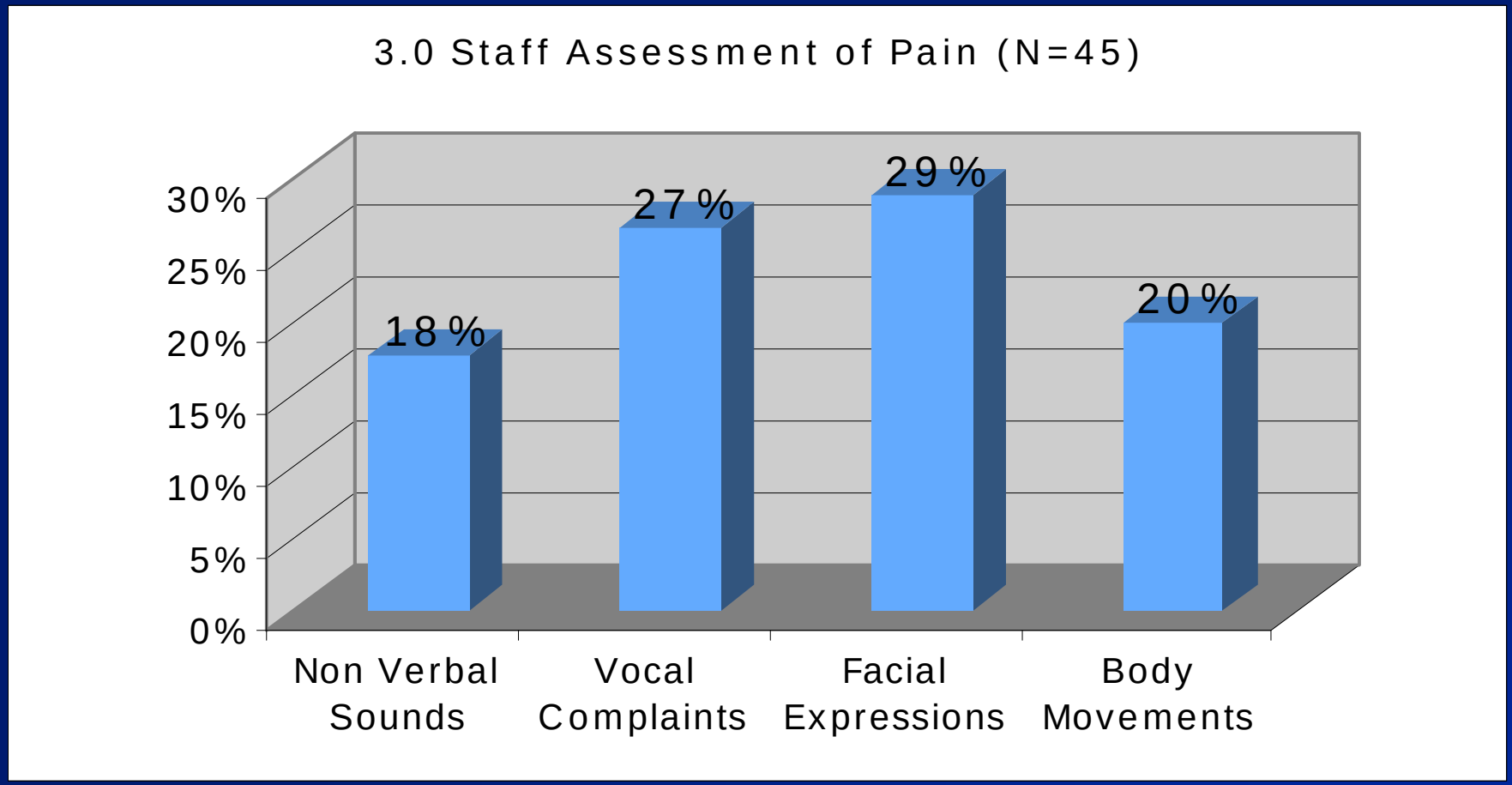
Temporal reliability

- Interview 24 hours later (different assessor also)
  - kappa = .9242 (.8837, .9647)

Agreement with MDS 2.0

  - kappa = .4812 (.3962, .5662)

Staff Assessment of Pain



43 % with any symptom

Other Sections with Important Changes

- **Pressure ulcer**
  - eliminated reverse staging
  - adds present on admit
- **Balance**
  - refocused on movement and transitions
- **Falls**
  - introduced type of injury
- **Bowel & bladder**
  - no longer rate catheter as continent
  - improved toileting program item

Other Sections with Important Changes

- **Activities of daily living – single response scale**
- **Goals of care & return to community items added**
- **Oral/dental item improved**
- **Swallowing item - checklist of observable signs and symptoms**
- **Restraints – separate bed and chair**

Overall, Nurses Judged MDS 3.0 Clinical Utility & Clarity Improved

- 85% rated MDS 3.0 as likely to help identify unrecognized problems
- 81% rated MDS 3.0 as more relevant than 2.0
- 84% reported that MDS 3.0 interview items improved their knowledge of the resident
- 85% rated MDS 3.0 questions as more clearly worded

Nurses also rated Validity High

- 89% rated MDS 3.0 as providing a more accurate report of resident characteristics than MDS 2.0
- 76% rated MDS 3.0 as better reflecting best clinical practice or standards

MDS 3.0 Took Less Time

MDS 3.0 Time **<** **MDS 2.0 Time**

- **Average time: 62 Min**
- **Average time: 112 Min**

Summary: MDS 3.0 Revisions are Based on

- Feedback from users
- Input from Experts
- Advances in assessment science
  - Improve clinical care in nursing home
  - Improve communication with providers
  - Improve ability to track care and patient progress across settings
- Testing of performance in NH populations

Summary: MDS 3.0 Revisions

- National testing showed increased resident voice and refined measures in MDS 3.0:
  - Increase measurement reliability
  - Increase measurement validity
  - Together these improve clinical detection and assessment accuracy
- Both Facility and Study Nurses from 71 NHs who used MDS 3.0 reported higher satisfaction due to:
  - Increased clinical relevance
  - Increased clarity
  - Increased knowledge about resident
- National testing showed reduced time to complete by 45%

RAND HEALTH

Questions?