NATIONAL QUALITY FORUM

National Voluntary Consensus Standards for Nursing Homes 2010

Measure Number/Title: NH-028-10: Consumer Assessment of Health Providers and Systems (CAHPS®) Nursing Home Survey: Family Member Instrument

Description: The CAHPS Nursing Home Survey: Family Member Instrument is a mail survey instrument to gather information on the experiences of family members of long stay (greater than 100 days) residents currently in nursing homes. The Centers for Medicare & Medicaid Services requested development of this questionnaire, which is intended to complement the CAHPS Nursing Home Survey: Long-Stay Resident Instrument and the Discharged resident Instrument. The Family Member Instrument asks respondents to report on their own experiences (not the resident's) with the nursing home and their perceptions of the quality of care provided to a family member living in a nursing home. The survey instrument provides nursing home level scores on 4 topics valued by patients and families: (1) Meeting Basic Needs: Help with Eating, Drinking, and Toileting; (2) Nurses/Aides' Kindness/ Respect Towards Resident; (3)Nursing Home Provides Information/Encourages Respondent Involvement; and (4) Nursing Home Staffing, Care of Belongings, and Cleanliness. In addition, the survey provides nursing home scores on 3 global items including an overall Rating of Care.

Numerator Statement: The following topics are measured for nursing homes from a family members persepctive:
Composite 1: Meeting Basic Needs – sum of applicable family member scores on 3 survey items (see codebook for points assigned to each response category) related to basic activities of daily living needs (help with eating, drinking, and toileting)
Composite 2: Nurses and Aides' Kindness and Respect towards Resident - sum of applicable family member scores on 5 survey items
Composite 3: How Well the Nursing Home Provides Information and Encourages Family Involvement - sum of applicable family member scores on 6 survey items
Composite 4: Nursing Home Staffing, Care of Belongings, and Cleanliness - sum of applicable family member scores on 7 survey items
Global Items:
Global Rating of care item: sum of family member scores on 0 to 10 scale
Global item whether ever unhappy with nursing home care: sum of family member scores on item (see codebook for points assigned to each response category)
Global item whether respondent would recommend nursing home: sum of family member scores on item (see codebook for points assigned to each response category).

Denominator Statement: The denominator is the total number of surveys for respondents that meet CAHPS completion standard and any applicable screener

Level of Analysis: Facility/Agency

Data Source: Survey: Patient, special or unique data

Measure developer: Agency for Healthcare Research and Quality

Type of Endorsement (full or time-limited): Full

Attachments: JASP_resident NHCAHPS; Integrated Nursing Home CAHPS Report; Codebook for Long Stay Nursing Home Residents; Nursing Home Final Report_17_Sept08

NATIONAL QUALITY FORUM

Measure Evaluation 4.1
December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the evaluation criteria are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized)**:** Complete all yellow highlighted areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the subcriteria (yellow highlighted areas).*

Steering Committee: Complete all pink highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** NH-028-10 **NQF Project:** Nursing Homes 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Consumer Assessment of Health Providers and Systems (CAHPS®) Nursing Home Survey: Family Member Instrument
De.2 **Brief description of measure***:* The CAHPS Nursing Home Survey: Family Member Instrument is a mail survey instrument to gather information on the experiences of family members of long stay (greater than 100 days) residents currently in nursing homes. The Centers for Medicare & Medicaid Services requested development of this questionnaire, which is intended to complement the CAHPS Nursing Home Survey: Long-Stay Resident Instrument and the Discharged resident Instrument. The Family Member Instrument asks respondents to report on their own experiences (not the resident's) with the nursing home and their perceptions of the quality of care provided to a family member living in a nursing home. The survey instrument provides nursing home level scores on 4 topics valued by patients and families: (1) Meeting Basic Needs: Help with Eating, Drinking, and Toileting; (2) Nurses/Aides´ Kindness/ Respect Towards Resident; (3)Nursing Home Provides Information/Encourages Respondent Involvement; and (4) Nursing Home Staffing, Care of Belongings, and Cleanliness. In addition, the survey provides nursing home scores on 3 global items including an overall Rating of Care.
1.1-2 **Type of Measure:** Patient experience De.3 **If included in a composite or paired with another measure, please identify composite or paired measure**
De.4 **National Priority Partners Priority Area:** Patient and family engagement De.5 **IOM Quality Domain**: Patient-centered De.6 **Consumer Care Need:**

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	**NQF Staff**
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a*	**A** Y☐

measure steward agreement even if measures are made publicly and freely available. **A.1 Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)?** Yes **A.2 Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** **A.3 Measure Steward Agreement:** Government entity and in the public domain - no agreement necessary **A.4 Measure Steward Agreement attached:**	**N**☐
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	**B** **Y**☐ **N**☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** Public reporting, Internal quality improvement	**C** **Y**☐ **N**☐
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. **D.1Testing:** Yes, fully developed and tested **D.2 Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** Yes	**D** **Y**☐ **N**☐
(for NQF staff use**) Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** **Y**☐ **N**☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. *Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.* (evaluation criteria) **1a. High Impact**	Eval Rating
(for NQF staff use) Specific NPP goal:	
1a.1 Demonstrated High Impact Aspect of Healthcare: Patient/societal consequences of poor quality **1a.2** **1a.3 Summary of Evidence of High Impact:** According to the 2004 National Nursing Home Survey (NNHS), there were approximately 1.5 million nursing home residents in 16,100 nursing home facilities (Jones et al, 2009). They are a population with significant limitations in activities of daily living (ADLs) with 51% receiving assistance with all 5 ADLs (bathing, dressing, toileting, transferring or eating)and less than 3% receiving no ADL help (Jones et al 2009); about 69% have cognitive impairment as measured by the Cognitive Performance Scale (CMS 2008). The National Health Expenditures Accounts(CMS, 2009) estimate that nursing home costs totaled $131 billion in 2008. With the passage of the Omnibus Reconciliation Act of 1987 (OBRA'87) Congress responded to growing concerns about the quality of care that nursing home residents received by requiring reforms in the federal certification and oversight of nursing homes. OBRA'87 shifted evaluations of health care quality from a	**1a** **C**☐ **P**☐ **M**☐ **N**☐

Comment [KP1]: 1a. The measure focus addresses:
- a specific national health goal/priority identified by NQF's National Priorities Partners; OR
- a demonstrated high impact aspect of healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use (current and/or future), severity of illness, and patient/societal consequences of poor quality).

focus on structure, and process criteria to clinical outcomes, resident satisfaction and quality of life. Since OBRA'87 implementation, GAO (2005; 2007)has continued to investigate quality of care in nursing homes and quality oversight activities of CMS and the states.
Concurrent with changes from OBRA'87 implementation, a radical rethinking of the long term care system known as "culture change" began more than a decade ago. Culture change refers to the transformation of nursing homes from an "acute care" model to a consumer-directed model. Common themes of changes include: autonomy in personal choices for the residents, improved communication between residents and staff,and more homelike environments (www.pioneernetwork.net). The Pioneer Network estimates that 5% of nursing homes have fully adopted culture change (www.pioneernetwork.net). Resident/Patient Experience surveys are one tool for a nursing home to use to become more resident-centered. Surveying family members is a very important source of feedback for nursing home residents who cannot respond independently to a survey(for example, residents with advanced dementia). The family also can often add information to the resident's viewpoint. The Institute of Medicine (2010) recently updated its conceptual framework for categorizing health care quality and disparities measurement to add family-centeredness to patient-centeredness. The National Priorities Partnership (http://www.nationalprioritiespartnership.org/PriorityDetails.aspx?id=596) also includes patient and family engagement as one of its priorities.

1a.4 Citations for Evidence of High Impact: Jones, A. L., Dwyer, L.L. , Bercovitz, A.R., Strahan, G. The National Nursing Home Survey: 2004 Overview. National Center for Health Statistics. Vital Health Stat. 13(167). 2009

CMS, Nursing Home Data Compendium, 2008 edition.
CMS national Health Expenditure Data is at http://www.cms.gov/NationalHealthExpendData/

GAO (Dec. 2005). "Despite increased oversight, challenges remainin ensuring high-quality care and resident safety" www.gao.gov/cgi-bin/getrpt?GAO-06-117.

GAO (May 2007). "Continued attention is needed to improve quality of care in small but significant share of homes." www.gao.gov/cgi-bin/getrpt?GAO-07-794T.

Institute of Medicine Committee on Future Directions for the National Healthcare Quality and Disparities Reports; Cheryl Ulmer, Michelle Bruno, and Sheila Burke, Editors; Future Directions for the National Healthcare Quality and Disparities Reports. Washington, DC: National Academy Press, 2010

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: The goal would be to use this family member survey as feedback to transform nursing home care to be resident-directed/centered and achieve the highest quality of life and quality of care for this vulnerable nursing home population.

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
The 2008 National Ombudsmen Reporting System (NORS) data showed that the top complaint of nursing home residents and their families, eliciting some 14,329 complaints to ombudsmen, was failing to respond to requests for assistance. The first composite, meeting basic needs, covers the top complaint identified by ombudsmen, indicating a critical need to assess how well and how poorly a nursing home provides basic care. Specific complaints relating to these items include lack of assistance with toileting which had 3,404 complaints; lack of assistance with drinking which had 2,899 complaints; and lack of assistance with eating which had 1,529 complaints (NORS, 2008). Similarly, most of the other negative items were also major sources of complaints. While no specific complaint used the word rude, complaints relating to dignity, respect and staff attitudes totaled 9,075. Fear of reprisals totaled 687—which may not seem high, but given the research indicating that people seldom complain about fear of reprisals, it suggests a significant issue. Finally, loss of laundry was mentioned 1,771 times in 2008. These common complaints are covered in the family member survey instrument.

1b
C ☐
P ☐
M ☐
N ☐

Comment [KP2]: 1b. Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating considerable variation, or overall poor performance, in the quality of care across providers and/or population groups (disparities in care).

Comment [k3]: 1 Examples of data on opportunity for improvement include, but are not limited to: prior studies, epidemiologic data, measure data from pilot testing or implementation. If data are not available, the measure focus is systematically assessed (e.g., expert panel rating) and judged to be a quality problem.

Under contract with CMS, states conduct nursing home inspections, known as surveys, to assess compliance with federal quality and safety requirements, including requriements for resident rights and quality of life. According to the CMS Nursing Home Compare website, the US average number of nursing home deficiencies issued as of March 2010 was 8; however the range of deficiencies by state was 0 to 68.

1b.3 Citations for data on performance gap:
1. National Ombudsmen Reporting System (NORS,2008). Top 20 complaints by category for nursing facilities (FFY 1996-2008). 2008 National Ombudsman Reporting System Data Tables (Unlettered Tables in Appendix B). Retrieved on December 31, 2009 from http://www.aoa.gov/AoARoot/AoA_Programs/Elder_Rights/Ombudsman/National_State_Data/2008/Index.aspx.

2. CMS Nursing Home Compare website contains information on U.S. average number of deficiency citations at www.medicare.gov/NHCompare

1b.4 Summary of Data on disparities by population group:
not available

1b.5 Citations for data on Disparities:
not available

1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*)**:** For consumer satisfaction/experience data to be useful to nursing homes (i.e., know what areas need improvement and which have priority), surveys should measure what is important to family members and residents. Survey data could also be used by consumers to help select higher quality nursing homes.

Some research indicates that higher family and resident satisfaction is associated with better resident clinical outcomes.

1c.2-3. Type of Evidence: Observational study, Expert opinion

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:**
Carefully developed patient experience surveys can inform nursing home providers about areas that need improvement particularly in areas that residents and families consider important.(see section 3a.6 for focus group results on what is important to consumers). These survey items complement the data nursing homes may currently collect to support improvements in internal customer services and quality related activities. Surveys can also be important to consumers for selecting nursing homes; however, surveys that have substantial ceiling effects may make it difficult to distinguish significant differences among nursing homes. This family member instrument had a fairly low percent at the ceiling for 3 out of 4 composites (9% to 15% for Composites 2, 3 and 4) while Composite 1 had 64% at ceiling because of limited degress of freedom with a Yes/No scale-- see Table 9 in AIR Final Report on page 24.

Two separate unpublished studies by Castle (personal communication, April 2010) indicate that higher family satisfaction is associated with fewer nursing home deficiency citations and clinical outcomes (less restraints and less depression).

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom***):**
ungraded

1c.6 Method for rating evidence: ungraded

1c.7 Summary of Controversy/Contradictory Evidence: none identified

1c.8 Citations for Evidence (*other than guidelines***):** Nicholas Castle, Ph.D., University of Pittsburgh

1c
C☐
P☐
M☐
N☐

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

Comment [k4]: 1c. The measure focus is:
•an outcome (e.g., morbidity, mortality, function, health-related quality of life) that is relevant to, or associated with, a national health goal/priority, the condition, population, and/or care being addressed;
OR
•if an intermediate outcome, process, structure, etc., there is **evidence** that supports the specific measure focus as follows:
oIntermediate outcome – evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
oProcess – evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and
if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
oStructure – evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
oPatient experience – evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
oAccess – evidence that an association exists between access to a health service and the outcomes of, or experience with, care. ... [1]

Comment [k5]: 4 Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multi-step process, the step with the greatest effect on the desired outcome should be selected as the focus of measurement. For example, although assessment of immunization status and recommending immunization are necessary steps, they are not sufficient to achieve the desired impact on health status – patients must be vaccinated to achieve immunity. This does not preclude consideration of measures of preventive screening interventions where there is a strong link with desired outcomes (e.g., mammography) or measures for multiple care processes that affect a single outcome.

Comment [k6]: 3 The strength of the body of evidence for the specific measure focus should be systematically assessed and rated (e.g., USPSTF grading system http://www.ahrq.gov/clinic/uspstf07/methods/benefit.htm). If the USPSTF grading system was not used, the grading system is explained including how it relates to the USPSTF grades or why it does not. However, evidence is not limited to quantitative studies and the best type of evidence depends upon the question being studied (e.g., randomized controlled trials appropriate for studying drug efficacy are not well suited for complex system changes). When qualitative studies are used, appropriate qualitative research criteria are used to judge the strength of the evidence.

(personal communication, April 2010), unpublished research from 2 study samples. (1) a sample of 6000 family members in 300 nursing homes (1200 units); and (2) a sample of 180 nursing homes with family, resident, and staff statisfaction surveys. 1c.9 **Quote the Specific guideline recommendation (***including guideline number and/or page number***):** not applicable 1c.10 **Clinical Practice Guideline Citation:** not applicable 1c.11 **National Guideline Clearinghouse or other URL:** not applicable 1c.12 **Rating of strength of recommendation (***also provide narrative description of the rating and by whom***):** not applicable 1c.13 **Method for rating strength of recommendation** (*If different from* USPSTF system*, also describe rating and how it relates to USPSTF)***:** not applicable 1c.14 **Rationale for using this guideline over others:** not applicable	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report***, met?** **Rationale:**	**1** **Y☐** **N☐**
2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	Eval Rating
2a. MEASURE SPECIFICATIONS	
S.1 **Do you have a web page where current detailed measure specifications can be obtained?** S.2 **If yes, provide web page URL:** **2a. Precisely Specified**	
2a.1 **Numerator Statement (***Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome***):** The following topics are measured for nursing homes from a family members persepctive: Composite 1: Meeting Basic Needs – sum of applicable family member scores on 3 survey items (see codebook for points assigned to each response category) related to basic activities of daily living needs (help with eating, drinking, and toileting) Composite 2: Nurses and Aides' Kindness and Respect towards Resident - sum of applicable family member scores on 5 survey items Composite 3: How Well the Nursing Home Provides Information and Encourages Family Involvement - sum of applicable family member scores on 6 survey items Composite 4: Nursing Home Staffing, Care of Belongings, and Cleanliness - sum of applicable family member scores on 7 survey items Global Items: Global Rating of care item: sum of family member scores on 0 to 10 scale Global item whether ever unhappy with nursing home care: sum of family member scores on item (see codebook for points assigned to each response category) Global item whether respondent would recommend nursing home: sum of family member scores on item (see codebook for points assigned to each response category).	**2a-specs** **C☐** **P☐** **M☐** **N☐**

Comment [k7]: USPSTF grading system **http://www.ahrq.gov/clinic/uspstf/grades.htm**: **A -** The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B -** The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C -** The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D -** The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I -** The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

Comment [KP8]: 2a. The measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability. The required data elements are of high quality as defined by NQF's Health Information Technology Expert Panel (HITEP) .

2a.2 Numerator Time Window (*The time period in which cases are eligible for inclusion in the numerator***):**
last six months

2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions***):**
Composite 1:3 survey items Q17, Q19, Q21
Composite 2: 5 survey items Q12, Q13, Q14, Q15, Q24
Composite 3: 6 survey items Q26, Q27, Q28, Q35, Q37, Q42
Composite 4: 7 survey items Q11, Q22, Q29, Q30, Q31, Q32, Q33, Q40
Global items: 3 survey items Q34, Q38, Q39

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured***):**
The denominator is the total number of surveys for respondents that meet CAHPS completion standard and any applicable screener (discussed in details below).

2a.5 Target population gender: Female, Male
2a.6 Target population age range: 18 and older

2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator***):**
last six months

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions***):**
Composite 1: Meeting Basic Needs:
Q17: the number of surveys completed by all those who responded "yes" to screener Q16
Q19: the number of surveys completed by all those who responded "yes" to screener Q18
Q21: the number of surveys completed by all those who responded "yes" to screener Q20
Composite 2: Nurses and Aides' Kindness and Respect towards Resident:
the denominator is the total number of completed surveys for 4 out of 5 questions in this composite excluding Q24; for Q24, its denominator is the number of surveys completed by all those who responded "yes" to screener Q23
Composite 3: How Well the Nursing Home Provides Information and Encourages Family Involvement:
the denominator is the total number of completed surveys for 2 out of 6 questions (Q27 and Q28) in this composite excluding these questions:
Q26: the number of surveys completed by all those who responded "yes" to screener Q25
Q35: the number of surveys completed by all those who responded "yes" to screener Q34
Q37: the number of surveys completed by all those who responded "yes" to screener Q36
Q42: the number of surveys completed by all those who responded "yes" to screener Q41
Composite 4: Nursing Home Staffing, Care of Belongings, and Cleanliness:
the denominator is the total number of completed surveys for 6 out of 7 questions in this composite excluding Q33; for Q33, its denominator is the number of surveys completed by all those who responded "yes" to screener Q32
Global Items: for all 3 global items the denominator is the total number of completed surveys.

2a.9 Denominator Exclusions (*Brief text description of exclusions from the target population***):** We exclude respondents who are under age 18, those who did not visit the nursing home resident at least once in 6 months, those whose resident was discharged, and those with a resident who had been in the nursing home for less than 100 days. In addition, screener questions may reduce the denominator size – those questions with screeners are noted in 2a.8 above.

2a.10 Denominator Exclusion Details (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions***):**
Q43-respondents age
Q9 -number of times visited nursing home resident in last 6 month
Q2 & Q3 resident was discharged

Comment [k9]: 11 Risk factors that influence outcomes should not be specified as exclusions.
12 Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

Q5 resident in the nursing home for less than 30 days

2a.11 Stratification Details/Variables (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions***):**
not applicable

2a.12-13 Risk Adjustment Type: Case-mix adjustment

2a.14 Risk Adjustment Methodology/Variables (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method***):**
The CAHPS team recommends four items to be case-mix adjusters for the CAHPS Nursing Home Family Survey: 1) respondent age, 2) respondent education, 3) whether the respondent believes the resident will permanently live in the nursing home, and 4) respondent's belief about whether the resident was capable of making decisions (See Table 10 on page 29 in AIR Final Report). Several additional items were considered as potential adjusters but were rejected for a variety of reasons. A full description of the risk adjustment process is available in the AIR Final Report on pages 26-33.

2a.15-17 Detailed risk model available Web page URL or attachment: Attachment Nursing Home Final Report (17 Sept 08).doc

2a.18-19 Type of Score: Non-weighted score/composite/scale
2a.20 Interpretation of Score:
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps***):**
SCORING AND PATIENT-MIX ADJUSTMENTS

1. Global rating and items
- Measured by family member's overall rating of the care at the nursing home on a scale of 0-10 (Q38)
- Measured by whether the family member was ever unhappy with the care their family member received at the nursing home on a Yes/No scale (Q34) Note: "No" represents better quality
- Measured by whether the family member would recommend the nursing home to others on a four-point scale: Definitely No, Probably No, Probably Yes, Definitely Yes (Q39)

2. Domains of care
1. Meeting Basic Needs – Help with Eating, Drinking and Toileting (Q17, Q19, & Q21)
2. Nurses and Aides' Kindness and Respect towards Resident (Q12, Q13, Q14, Q15, & Q24)
3. How Well the Nursing Home Provides Information and Encourages Family Involvement (Q26, Q27, Q28, Q53, Q37 & Q42)
4. Nursing Home Staffing, Care of Belongings, and Cleanliness (Q11, Q22, Q29, Q30, Q31, Q33 & Q40)

3. Production of nursing home scores – Global items

- Q38 Nursing home level ratings are presented using percentages for three-categories for the 0-10 scale question: 0-6, 7-8, and 9-10.
- Q39 Nursing home level scores are presented using percentages for the following three categories: definitely would recommend, probably would recommend, and definitely not or probably not recommend
- Q34 Nursing home level scores are presented using percentages for two categories (reverse coded): yes, happy with nursing home care in past 6 months; and no, not happy with nursing home care in past 6 months

4. Production of nursing home scores - Domain-level composites
There are four domain-level composites included in the Nursing Home Family Member Questionnaire: 1) Meeting Basic Needs – Help with Eating, Drinking and Toileting; 2) Nurses and Aides' Kindness and Respect towards Resident; 3) How Well the Nursing Home Provides Information and Encourages Family Involvement; 4) Nursing Home Staffing, Care of Belongings, and Cleanliness.

- Meeting Basic Needs – Help with Eating, Drinking and Toileting
This composite is produced by combining responses to three questions:
- Q17: Family member helped nursing home resident with eating. "Was it because the nurses or

aides either didn't help or made him or her wait too long?"
- Q19: Family member helped nursing home resident with drinking. "Was it because the nurses or aides either didn't help or made him or her wait too long?"
- Q21: Family member helped nursing home resident with toileting. "Was it because the nurses or aides either didn't help or made him or her wait too long?"

Respondents can answer "yes" or "no" to each. (note: "yes" represents lower quality" A nursing home's score on the "Meeting Basic Needs – Help with Eating, Drinking and Toileting" composite is the proportion of cases in each response category.

The steps to calculate a nursing home provider's composite score follow:

Step 1 – Calculate the proportion of cases in each response category for the first question:
P11 = Proportion of respondents who answered "yes"
P12 = Proportion of respondents who answered "no"

Follow the same steps for the second question:
P21 = Proportion of respondents who answered "yes"
P22 = Proportion of respondents who answered "no"

Follow the same steps for the third question:
P31 = Proportion of respondents who answered "yes"
P32 = Proportion of respondents who answered "no"

Step 2 – Combine responses from the questions to form the composite
Calculate the average proportion responding to each category across the questions in the composite. For example, in the "Meeting Basic Needs – Help with Eating, Drinking and Toileting" composite (three questions), calculations would be as follows:
PC1 = Composite proportion who said "yes" = (P11 + P21 + P31) / 3
PC2 = Composite proportion who said "no" = (P12 + P22 + P32) / 3

- Nurses and Aides' Kindness and Respect towards Resident

This composite is produced by combining responses to five questions:
- Q12: "In the last 6 months, how often did you see the nurses and aides treat your family member with courtesy and respect?"
- Q13: "In the last 6 months, how often did you see the nurses and aides treat your family member with kindness?"
- Q14: "In the last 6 months, how often did you feel that the nurses and aides really cared about your family member?"
- Q15: "In the last 6 months, did you ever see any nurses or aides be rude to your family member or any other resident?" *(Yes/no)
- Q24: "In the last 6 months, how often did the nurses and aides handle the situation in a way that you felt was appropriate?"

Respondents can answer "never," "sometimes," "usually," or "always" to each item except for Q15 where the response scale is Yes/No. The steps to calculate a nursing home's composite score for this domain are the following:

Step 1 – Calculate the proportion of cases in each response category for the first question (Q12):
P11 = Proportion of respondents who answered "never"
P12 = Proportion of respondents who answered "sometimes"
P13 = Proportion of respondents who answered "usually"
P14 = Proportion of respondents who answered "always"

Follow the same steps for the second (Q13), third (Q14), and fifth (Q26) questions:

For the fourth question (Q15) calculate the proportion of cases in each response category: note: "No" represents better quality

P41= Proportion of respondents who answered "yes"
P42= Proportion of respondents who answered "no"

Step 2 – Combine responses from the questions to form the composite
Calculate the average proportion responding to each category across the questions in the composite. For example, in the "Nurses and Aides' Kindness and Respect Towards Resident" composite (five questions), calculations would be as follows:

PC1 = Composite proportion who said "never" or "yes" = (P11 + P21 + P31 + P41 + P51) / 5 note: "yes" represents worse quality
PC2 = Composite proportion who said "sometimes" = (P12 + P22 + P32 + P52) / 5
PC3 = Composite proportion who said "usually" = (P13 + P23 + P33 + P53) / 5
PC4 = Composite proportion who said "always" or "no"= (P14 + P24 + P34 + P44 + P54) / 5 note: "No" represents better quality

Survey sponsors may choose an alternative to combine proportions of respondents who said "never" or "sometimes" or "yes" and compare with combined proportions of respondents who said "always" or "usually" or "no".

- How Well the Nursing Home Provides Information and Encourages Family Involvement

This composite is produced by combining responses to six questions:
- Q26: "In the last 6 months, how often did you get this information as soon as you wanted?"
- Q27: "In the last 6 months, how often did the nurses and aides explain things in a way that was easy for you to understand?"
- Q28: "In the last 6 months, did the nurses and aides ever try to discourage you from asking questions about your family member?" note: "No" represents better quality
- Q35: "In the last 6 months, did you ever stop yourself from talking to any nursing home staff about your concerns because you thought they would take it out on your family member?" note: "No" represents better quality
- Q37: "In the last 6 months, how often were you involved as much as you wanted to be in the decisions about your family member's care?"
- Q42: "In the last 6 months, how often did you get all the information you wanted from the nursing home about payments or expenses?"

Respondents to four of the above questions can answer "never," "sometimes," "usually," or "always" to each. Respondents can answer "yes" or "no" to two of the above questions (Q31 and Q43), where "no" indicates better quality. The steps to calculate a nursing home's composite score for this domain are similar to calculations for Composite 2: "Nurses and Aides' Kindness and Respect Towards Resident", except that in Step 2, each composite proportion category would be divided by six (the total number of items).

- Nursing Home Staffing, Care of Belongings, and Cleanliness

This composite is produced by combining responses to seven questions:
- Q11: "In the last 6 months, how often were you able to find a nurse or aide when you wanted one?"
- Q22: "In the last 6 months, how often did your family member look and smell clean?"
- Q29: "In the last 6 months, how often did your family member's room look and smell clean?"
- Q30: "In the last 6 months, how often did the public areas of the nursing home look and smell clean?"
- Q37: "Personal medical belongings are things like hearing aids, glasses, and dentures. In the last 6 months, how often were your family member's personal medical belongings damaged or lost?" note: "Never" represents better quality
- Q39: "In the last 6 months, when your family member used the laundry service, how often were clothes damaged or lost?" note: "Never" represents better quality
- Q51: "In the last 6 months, how often did you feel there were enough nurses and aides in this nursing home?"

Respondents to three of the above questions can answer "never," "sometimes," "usually," or "always" to

each. Respondents to two of the above questions (Q37 and Q39) can answer "never", "once", or "two or more times" to each, where "Never" represents better quality. The steps to calculate a nursing home's composite score for this domain are similar to calculations for Composite 2: "Nurses and Aides' Kindness and Respect Towards Resident", except that in Step 2, each composite proportion category would be divided by seven (the total number of items).

Risk adjustment algorithm is provided as attachment in Additonal Information section at Ad.11

2a.22 Describe the method for discriminating performance *(e.g., significance testing)*:
For statistical significance for each composite or global item, we used a t-test comparing each nursing home mean to the mean of all the nursing home means.

2a.23 Sampling (Survey) Methodology *If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate)*:
Sampling Frame Elements: An eligible sample member is the person listed by the nursing home as the responsible person for a resident who has resided at the nursing home for at least 30 consecutive days. Eligible sample members can include family, friends, guardians, people with medical power of attorney for the resident, and attorneys. This survey is designed for adults only (18 and older). If a resident listed more than one responsible party, the respondent should be randomly selected. If the same responsible party is listed for more than one resident, the resident for whom the responsible party responds should be randomly selected. If there is more than one responsible party listed for a resident, randomly select one of them.
The sampling frame should include: Name of responsible party, Address, Telephone number; Resident/patient name; date of birth, gender, whether the responsible party was the power of attorney; admission date; and whether the resident is in a dementia unit.
Drawing the Sample: Based on the CAHPS grantees' experiences with the field tests of this instrument, we recommend the following:

- For facilities with up to 150 eligible patients, use all patients (a census) from each facility.
- For facilities with more than 150 eligible patients, draw a systematic random sample of 150 patients from each facility. If you anticipate that poor contact information (addresses and telephone numbers) will decrease the number of questionnaires that reach the sampled individuals, you may need to start with a larger sample.

Data Collection Protocol: Recommended Protocol for Mail with Optional Telephone Followup: The following guidance builds upon the grantees' experiences fielding CAHPS and other surveys, as well as their specific experience with the field test of the Family Member Instrument. The CAHPS Team recommends using one of the following two protocols for data collection:

- Two (2) mailings of the survey with a reminder postcard/letter prior to the 2nd mailing followed by telephone contact for those family members who have not responded to the mailed surveys.
- Two (2) mailings of the survey with a reminder postcard/letter prior to the 2nd mailing without the telephone followup.

Once the vendor has initiated the data collection process, it is important to follow the protocol through to completion. Even if you achieve the minimum response rate of 50 percent, continue with the survey administration protocol to achieve the highest response rate possible.
In the field test of this instrument, the CAHPS Consortium tested elements of this survey administration protocol by mail with telephone followup. A response rate of 66 percent was achieved through a combination of the following:

- An initial mailing of the questionnaire with a cover letter and return postage-paid envelope : 42% response rate
- A second mailing of the questionnaire 2 weeks after the reminder: 14 %
- Computer-assisted telephone interviews (CATI) for non-respondents 2 weeks after the second mailing of the questionnaire: 10 %

Minimum sample size:

The number of subjects needed for each composite to reach a reliability of 0.70 (if the goal is public reporting for reliable comparison purposes) was calculated with the Spearman-Brown Prediction formula using the average number of respondents per nursing home. This number was then adjusted by the lowest proportion that are eligible for any question in the composites (so that reliability is achieved on all scales). Based on the pilot test of the family member survey, Q23 (waited too long for help with toileting) had the

lowest proportion (25%) who were eligible to respond (based on the screener Q22). So the number needed to reach 0.70 reliability for the composite Meeting Basic Needs was (31/0.25) or 124. If necessary this data could be accumulated over time to achieve sufficient sample size. The other 3 composites require smaller sample numbers:
Composite 2: Nurses/Aides' Kindness/Respect Towards Resident: (6.3/.26)= 24.2 minimum recommended
Composite 3: How Well the Nursing Home Provides Information and Encourages Family Involvement: (6.4/.32)= 20 minimum recommended
Composite 4: Nursing Home Staffing, Care of Belongings and Cleanliness: (10.8/.765)= 14 minimum recommended

If the goal is to use survey data only for quality improvement purposes, a smaller number of completes may be used.

2a.24 Data Source (*Check the source(s) for which the measure is specified and tested)*
Survey: Patient, Special or unique data

2a.25 Data source/data collection instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.***):**
CAHPS Nursing Home Survey: Family Member Instrument

2a.26-28 Data source/data collection instrument reference web page URL or attachment: URL
https://www.cahps.ahrq.gov/content/products/NH/652_NHFamily_Eng.pdf

2a.29-31 Data dictionary/code table web page URL or attachment: Attachment CODEBOOK FAMILY MEMBER NURSING HOME SURVEY final 5_7_10.doc

2a.32-35 Level of Measurement/Analysis (*Check the level(s) for which the measure is specified and tested)*
Facility/Agency

2a.36-37 Care Settings (*Check the setting(s) for which the measure is specified and tested)*
Nursing home (NH) /Skilled Nursing Facility (SNF)

2a.38-41 Clinical Services (*Healthcare services being measured, check all that apply*)
Clinicians: Pharmacist, Other nurse aides

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample *(description of data/sample and size)***:** With the assistance of the Texas State Long Term Care Ombudsman, the CAHPS Team conducted a field test at 15 nursing homes in Texas between October 2006 and January 2007. Of the 1,444 family members in the sample with addresses, 885 completed the survey for a response rate of 66%.

2b.2 Analytic Method *(type of reliability & rationale, method for testing)***:**
To look at reliability, internal consistency reliability (alpha) was estimated. This is a measure of how well the items in a composite hang together. Composites should have an alpha of 0.70 or greater to be considered reliable. Additionally, we looked at nursing-home (NH)-level reliability, or inter-unit reliability (IUR). This statistic represents a transformation of the F-statistic for testing differences among agencies on an item or composite (IUR = (F-1)/F). IUR can be interpreted as the fraction of the variation among facility scores that is due to real differences, rather than due to chance. If the IUR is higher, the ability of the item or composite to discriminate across facilities is greater. An IUR > 0.70 is considered to indicate a high level of discriminant ability for an item or composite. As the IUR gets smaller, you need a larger sample in order to reliably discriminate across facilities.

2b.3 Testing Results *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)***:**
All four composites had a Cronbach's alpha greater than 0.7 which indicates that the scores would provide

2b
C☐ P☐ M☐ N☐

Comment [KP10]: 2b. Reliability testing demonstrates the measure results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period.

Comment [k11]: 8 Examples of reliability testing include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing may address the data items or final measure score.

reliable data. The alphas were: Meeting Basic Needs = 0.90; Nurses and Aides' Kindness and Respect towards Family Members=0.88; How Well the NH Provides Information and Encourages Family Involvement = 0.78; and NH Staffing, Care of Belongings and Cleanliness = 0.79. For more detail see pages 22-25 and Tables 7, 8 and 9 in the AIR Final report.
Three composites (except Meeting Basic Needs) had NHr (or IUR-"inter-unit reliability") greater than 0.7: Nurse & Aides Kindness/Respect Towards Resident was 0.83; Nursing home Provides Information/Encourages respondent Involvement was 0.85 and Nursing Home Staffing, Care of Belongings and Cleanliness was 0.89. Although the observed facility-level reliability of the Meeting Basic Needs composite is not as high (NHr= 0.48) as we would like, it will be able to discriminate across nursing homes, given a sufficient number of respondents per facility. Because this composite is made up of three items that were appropriately skipped by a large number of respondents, it has a high percentage of missing data. (See Table 7 and Table 8 on pages 22-24 in AIR Final report).

These final 4 domains were a balance of theory (original intent of items and composites), statistical evidence of reliability and validity (item-level and nursing home level, factor analyses) and stakeholder perspectives. Ten individual items were recommended to be dropped due to low statistical values (see Table 3 in AIR Final Report). Two items ("Nurses & Aides Discourage Questions" and "Medical Belongings Lost") were retained despite marginal measurement characteristics because of stakeholder interest and because they scaled well with their respective composite (see Table 3 in AIR Final Report).

2c. Validity testing

2c.1 **Data/sample** *(description of data/sample and size)*: Data are from the field test at 15 nursing homes in Texas between October 2006 and January 2007 (n=885)

2c.2 **Analytic Method** *(type of validity & rationale, method for testing)*:
We examined the correlation of each of the composites with the global ratings as a measure of criterion validity. We also used scaling success to summarize the discriminant validity of the composite measures, or the degree to which each item correlates more highly with its own scale than it does with competing scales. The target for scaling success should be about 100%.

2c.3 **Testing Results** *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
All four composites demonstrate sufficient criterion validity, as evidenced by their relatively high correlations (> 0.30) with the three global measures (see Table 7 on page 23 of AIR Final Report). Except for the composite on NH Staffing, Care of Belongings and Cleanliness, the scaling success was 100%; the 4th composite had a scaling success of 86%. The `statistical results taken together for this last composite indicate that the composite has reliable scores but overlaps in meaning with some content in the other composites. This finding is to be expected given that the composite is a general indicator of nursing home quality and actually indicative of the composite's validity as a more general measure. (see Tables 8 and 9 and pages 24-25 in AIR Final report for more detail)

2c
C☐
P☐
M☐
N☐

2d. Exclusions Justified

2d.1 **Summary of Evidence supporting exclusion(s):**
There are several exclusions for the family member survey. The first one is family members of residents who have resided in the nursing home for less than 30 consecutive days; another is family members who are less than 18 years old. There are two reasons for the first exclusion. The CAHPS team believed that a minimum of 30 days was needed for a family member to obtain a relatively stable opinion about facility care for a resident who will be a long stay resident. Another reason is that AHRQ created a separate mail survey for short stay residents whose opinion can be directly obtained because on average they have less cognitive impairment than long stay residents. For family members who are under 18 years old, the CAHPS team decided they would not accept those under age 18 as responsible adults since 18 and older is generally regarded as adult.

2d.2 **Citations for Evidence:**
Expert opinion, cognitive testing and sample frame development for field test

2d
C☐
P☐
M☐
N☐
NA☐

Comment [KP12]: 2c. Validity testing demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed.

Comment [k13]: 9 Examples of validity testing include, but are not limited to: determining if measure scores adequately distinguish between providers known to have good or poor quality assessed by another valid method; correlation of measure scores with another valid indicator of quality for the specific topic; ability of measure scores to predict scores on some other related valid measure; content validity for multi-item scales/tests. Face validity is a subjective assessment by experts of whether the measure reflects the quality of care (e.g., whether the proportion of patients with BP < 140/90 is a marker of quality). If face validity is the only validity addressed, it is systematically assessed (e.g., ratings by relevant stakeholders) and the measure is judged to represent quality care for the specific topic and that the measure focus is the most important aspect of quality for the specific topic.

Comment [KP14]: 2d. Clinically necessary measure exclusions are identified and must be:
•supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion;
AND
•a clinically appropriate exception (e.g., contraindication) to eligibility for the measure focus;
AND
•precisely defined and specified:
–if there is substantial variability in exclusions across providers, the measure is specified so that exclusions are computable and the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion);
if patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that it strongly impacts performance on the measure and the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

Comment [k15]: 10 Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, sensitivity analyses with and without the exclusion, and variability of exclusions across providers.

2d.3 Data/sample *(description of data/sample and size)*: Data for sample frame development are from the field test at 15 nursing homes in Texas between October 2006 and January 2007.

2d.4 Analytic Method *(type analysis & rationale)*:
Cognitive testing included participants with family members who were discharged from a nursing home. Also, developing the sample frame for the field test in Texas gave the team information about different types of persons listed as responsible parties.

2d.5 Testing Results *(e.g., frequency, variability, sensitivity analyses)*:
Cognitive testing showed that participants with family members who were discharged were a different audience from those who had family members currently residing in the nursing home.
During development of the sample frame of the list of eligible responsible parties from the nursing home, the team excluded or re-designated the eligible responsible party when, in their judgment, the named party would have been uninformed or biased about the nursing home's performance. This included: (1).31 cases, representing 2.1% of the responsible parties received from the nursing homes, where the responsible party was a nursing home resident (including self) or a nursing home staff member or a trust fund or another institution. All of these were excluded. (2). 28 cases (1.9%), where the responsible party listed was the resident or the nursing home, but there was another, eligible contact listed as well. In these cases, the other contact was used instead of excluding the case. (3). 44 cases (3.0%), where residents had more than one eligible responsible party listed. This does not include the other contacts most residents had listed; only cases in which multiple persons were listed as responsible parties for the same resident. In these cases, one person was randomly selected to participate in the survey. (4). Another 17 eligible sample members (1.2%) listed as the responsible party for more than one resident. For these cases, the team randomly selected a resident for them to respond about. (5). 12 cases (0.8%), where an eligible responsible party was listed without contact information, but another person listed as a contact (not a responsible party) had adequate contact information. In these cases, they used the alternative with adequate contact information.
In addition to the above exclusions based on administrative data, the questionnaire included several screening questions that excluded other cases. It screened out eligible sample members who had visited the focal resident less than once in the last six months, and whose focal resident had been discharged from the focal nursing home at the time of the survey. Participants with family members who were discharged were excluded, because it became apparent during cognitive testing, that these were two different audiences and a separate instrument was needed to examine experiences of people who had a nursing home resident recently discharged or transferred from a nursing home.

2e. Risk Adjustment for Outcomes/ Resource Use Measures

2e.1 Data/sample *(description of data/sample and size)*: Data are from the field test at 15 nursing homes in Texas between October 2006 and January 2007 (n=885)

2e.2 Analytic Method *(type of risk adjustment, analysis, & rationale)*:
Four items are recommended as case-mix adjusters for the CAHPS Nursing Home Family Survey: 1) respondent age, 2) respondent education, 3) whether the respondent believes the resident will permanently live in the nursing home, and 4) respondent's belief about whether the resident was capable of making decisions (See Table 10 on page 29 in AIR Final Report).

2e.3 Testing Results *(risk model performance metrics)*:
The CAHPS team recommends four items to be case-mix adjusters for the CAHPS Nursing Home Family Survey: 1) respondent age, 2) respondent education, 3) whether the respondent believes the resident will permanently live in the nursing home, and 4) respondent's belief about whether the resident was capable of making decisions (See Table 10 on page 29 in AIR Final Report). Several additional items were considered as potential adjusters but were rejected for a variety of reasons. A full description of the risk adjustment process is available in the AIR Final Report on pages 26-33.

2e.4 If outcome or resource use measure is not risk adjusted, provide rationale:

2e
C☐
P☐
M☐
N☐
NA☐

2f. Identification of Meaningful Differences in Performance

2f
C☐
P☐

2f.1 Data/sample from Testing or Current Use *(description of data/sample and size)*: Data are from the

Comment [KP16]: 2e. For outcome measures and other measures (e.g., resource use) when indicated:
•an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present at start of care;[Error! Bookmark not defined.] OR rationale/data support no risk adjustment.

Comment [k17]: 13 Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care such as race, socioeconomic status, gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than adjusting out differences.

Comment [KP18]: 2f. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful differences in performance.

field test at 15 nursing homes between October 2006 and January 2007 (n=885). **2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*: For statistical significance, we used t-test comparing each nursing home mean to the mean of all the nursing home means. **2f.3 Provide Measure Scores from Testing or Current Use** *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)*: The means and standard deviations (SD) for the composites are: Composite 1: Meeting Basic Needs-- mean = 73.5 (39.3) Composite 2: Nurses and Aides' Kindness and Respect towards Family Members-- mean = 84.8 (19.9) Composite 3: How Well the NH Provides Info and Encourages Family Involvement – mean= 87.4 (17.0) Composite 4: NH Staffing, Care of Belongings, and Cleanliness – mean = 80.5 (17.1) Frequencies from the pilot test are avaialbe for all pilot survey questions in Appendix G of AIR Final Report	M☐ N☐
2g. Comparability of Multiple Data Sources/Methods **2g.1 Data/sample** *(description of data/sample and size)*: The Health Quality Council of Alberta conducted its own field test in 14 nursing homes and supplied data to AIR for psychometric analyses **2g.2 Analytic Method** *(type of analysis & rationale)*: **2g.3 Testing Results** *(e.g., correlation statistics, comparison of rankings)*: available on request	**2g** C☐ P☐ M☐ N☐ NA☐
2h. Disparities in Care **2h.1 If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: not applicable - measure is not stratified **2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:**	**2h** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, **met?** **Rationale:**	*2* C☐ P☐ M☐ N☐
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	Eval Ratin g
3a. Meaningful, Understandable, and Useful Information **3a.1 Current Use:** In use **3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large)** (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not*	**3a** C☐ P☐ M☐ N☐

Comment [k19]: 14 With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74% v. 75%) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v. $5,025) is practically meaningful. Measures with overall poor performance may not demonstrate much variability across providers.

Comment [KP20]: 2g. If multiple data sources/methods are allowed, there is demonstration they produce comparable results.

Comment [KP21]: 2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender);OR rationale/data justifies why stratification is not necessary or not feasible.

Comment [KP22]: 3a. Demonstration that information produced by the measure is meaningful, understandable, and useful to the intended audience(s) for both public reporting (e.g., focus group, cognitive testing) and informing quality improvement (e.g., quality improvement initiatives). An important outcome that may not have an identified improvement strategy still can be useful for informing quality improvement by identifying the need for and stimulating new approaches to improvement.

publicly reported, *state the plans to achieve public reporting within 3 years*)**:**
The Health Quality Council of Alberta, Canada, is using this survey for public reporting in aggregate -- see http://hqca.ca/index.php?id=130.

3a.3 **If used in other programs/initiatives (***If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):**
The Health Quality Council of Alberta, Canada, is using this survey for QI by providing site specific results back to nursing homes and comparing them to peers and norms. Also, this survey is included as one possible survey for nursing homes to use as part of Goal 7 (measuring Resident & Family Satisfaction) of the Advancing Excellence in America's Nursing Homes Campaign, of which more than 6400 U.S. nursing homes have joined (a home should pick three out of 8 possible goals).

Testing of Interpretability (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*)
3a.4 **Data/sample** *(description of data/sample and size)***:** Anyone with a family member in a nursing home was eligible to participate in the cognitive interviews; however persons were selected so as to assure variation in race, ethnicity, and education. The team conducted a total of 27 interviews in the first testing round in June 2005 and conducted another 27 interviews in the second round in June 2006.

3a.5 **Methods** *(e.g., focus group, survey, QI project)***:**
The formative research included focus groups, a call for measures, and reviewing literature. The goals of the focus groups were to: a) Understand participants' current experiences with nursing homes; b) Determine how participants' conceptualize good care; c) Determine the comparative salience and importance of the factors associated with good care; and d) Understand participants' potential uses of a nursing home quality report. In order to answer these questions, the three CAHPS grantees (Harvard University, AIR, and RAND, conducted 12 focus groups: two each in New York City and Phoenix, Arizona and four each in Palo Alto, California and Chapel Hill, North Carolina. With the exception of those in North Carolina, all focus group participants had already chosen a nursing home for a family member. Those participants in North Carolina were at the stage of considering moving a relative to a nursing home within a year. Detailed findings from the focus groups can be found in Appendix A, Focus Group Findings from AIR, RAND, and Harvard. AHRQ also published a call for measures in the Federal Register.

After reviewing the questionnaires and items received in response to that notice, and combined with information from a literature review and the focus groups, the team prepared a draft instrument for testing purposes. In the pilot version of the Nursing Home Family Member Survey, there were 12 negatively framed items (see Table 1 in paper posted at http://www.fcsm.gov/09papers/Frentzel_XI-C.pdf.) For the Nursing Home Family Survey, one purpose of using negative items has been to measure unique constructs that could not be framed as a positively written item. Because of the significant issue of staff failing to respond to requests for assistance in nursing homes, the CAHPS team developed items that would explore the issue of staff responsiveness on the most essential ADLs – eating, drinking fluids and toileting. For example, the survey asks a screener question, ?In the last 6 months, during any of your visits, did you help your family member with eating? and if yes, the survey then asks, ?Was it
because the nurses or aides either didn't help or made him or her wait too long?Two similar items to this on drinking and toileting are also asked. Had these items been written in a positive frame, e.g., ?How often did the nurses or aides help your family member with eating?, the items would capture a different construct. Similarly, the item, "in the last 6 months, did the nurses and aides ever try to discourage you from asking questions about your family member?" would not have the same meaning if positively framed, for example, "in the last 6 months, did the nurses and aides ever try to encourage you to ask questions about your family member?" When positively framed, it
measures a different construct that does not help to elucidate a potentially significant problem in the nursing home, i.e., that the patient population holds back from talking about a problem because of fear of reprisals.

In 2005 and 2006, the items in this draft instrument went through two rounds of cognitive testing with people who have family members in nursing homes. The cognitive interviews examined the following issues related to the draft questionnaires: 1. Content: Are the questions that are included in the survey important to consumers? Are consumers able to make judgments about the questions? 2. Comprehension and

Interpretation: Are the words, phrases, and questions easy to understand regardless of education level or knowledge of nursing home care? Are the questions interpreted as intended? 3. Recall: Are consumers able to recall the events asked about and to make judgments about them? 4. Navigation: Did the survey flow correctly? Were people skipped out of sections appropriately? The team prepared a formatted "cognitive testing" version of the instrument and a cognitive interviewing protocol for use by the interviewers. This protocol provided a listing of scripted probes that could be employed to provide insights into each respondent's cognitive processes as he or she read and answered the pilot items. It also included a series of general questions about the items, to allow the respondent to provide additional feedback about the items and to help assess the comprehensiveness of the instrument. A think aloud training exercise, with practice questions and a scripted response for the interviewers to use in modeling appropriate thinking aloud behaviors, was also included (for more detail see Appendix B: Draft Survey and First Cognitive Testing Protocol). Following the first round of cognitive tests, the team revised the items and prepared a protocol for the second round of cognitive testing (for more detail see Appendix D: Draft Survey and Second Cognitive Testing Protocol). This round tested the items as both self-administered items and as interviewer-administered items under the assumption that the final instrument would probably be administered by both mail and telephone. The instrument was revised again after the second round of cognitive testing. Trained cognitive interviewers conducted one-on-one, in-person interviews. Using a "think-aloud" approach for the interviews, the researchers asked the participant to read each question (or read the question verbally to the participant), provide a verbal response, and explain the reason for the response. They then followed up with probes after each of the questions to ask about specific potential problems with each item. Immediately after each interview, the interviewer wrote a summary of the participant's comments. A member of each grantee team reviewed these summaries and extracted common themes for each item. At the end of each round, the team met to review these themes and make recommendations to the larger family survey team. The larger team made decisions together regarding each of the items.

3a.6 Results *(qualitative and/or quantitative results and conclusions)*:
Focus Groups Summary:
The factors or components of nursing home quality identified by participants represent a mix of quality of life, quality of care, safety, and security.
Aspects of quality that were mentioned in the focus groups included: (1) Quality and the type of medical or physical care available at the nursing home; (2) Physical aspects of the nursing home facility, for example, cleanliness, security, appearance, size of rooms; (3) Quality of the nursing home staff, referring to whether they are trained or certified, and whether staffing numbers are adequate; (4) Quality of the nursing home staff, referring to whether the nursing home staff members are caring, approachable, and dedicated; (5) Quality and type of activities and opportunities for social engagement; (6) Adequate information-sharing and responsiveness to the family; (7) Respect for the resident and treating the resident as a "real person." In addition to the factors above, participants mentioned that they would like nursing homes to be required to post their "scores." Overall, participants reported that both the resident and family survey information would be useful in helping them choose a nursing home. However, a few participants did voice a concern that the resident surveys may not be as reliable because of cognitive constraints, fear of retaliation, and other personal characteristics. The family survey could help to address some of the weaknesses in the resident survey because people felt that family members would have useful observer information. On the other hand, people also noted some potential weaknesses associated with family information: "Family members might overlook some things depending on what their relationship was. They figure 'well this is better than having them somewhere else.'" Participants stated that they would use survey information in two main ways: before visiting a nursing home and, similarly, to narrow down the choices. (For more detail see Appendix A of AIR Final report)

Cognitive testing results summary:
In the first Round of cognitive testing, the most important issues considered and resolved were: 1) Asking questions about the respondent's experiences: We found that some of the questions we tested were actually proxy questions about the family member's experience. We had tried to ask them if they "observed" something, but in testing found that many of the items were things that visitors just didn't have experience with and didn't really see (basically the care processes). Details of this can be found in the question by question report. Solution: Delete proxy questions and only ask items in which the family member reports or rates a direct experience that they had. 2) Answering for all family members, not just self. Solution: Add language in the instrument and the cover letter asking R to please answer only about "yourself and your own personal experiences"; 3) The label: People had trouble with the date on the label.

It was confusing and not attended to correctly. Solution: Since we will now have separate instruments for current and discharged residents, we can remove the date from the label and add a time frame to the questions ("in the last 6 months"). Also, since this instrument focuses only on those still in a nursing home, we can eliminate the phrase "when you visited" from most questions.
In the second Round of cognitive testing, the most important issues considered and resolved were: 1) Q8: "family member capable of making decisions" - The item rationale needs to focus only on cognitive impairment and revise the question accordingly--[FINAL version] Is your family member able to make decisions about his or her own daily life, such as when to get up, what clothes to wear, and which activities to do? NSUA [based on MDS Problem, evaluates either or both family capability and stringent rules of nursing home]; 2) Q11/12: (courtesy & respect/kindness) - as always, "kindness" goes beyond "courtesy & respect" (which is expected). EF: In most cases we found similar findings, but… Respondents generally defined courtesy and respect as the same and had similar definitions for kindness. However, two respondents defined courtesy, respect, and kindness as the same. One theme that came up was that aides are paid to be courteous and respectful, but kindness is a personal attribute they bring to their job, "kindness you do on your own." For now, suggest either leaving it in or waiting to see the field test results to see if items if there is no added discrimination provided by the second item. Testing Decision: Keep items in as is; 3) Q20/21: (help toileting) - this question does not specify "in the nursing home" - 2 respondents said yes because they had to help on trips outside the NH (there was no staff there, so Q21 could not be answered) - 1 respondent understood "toileting" to be "on the toilet" and said "no" because they were now in Depends and not using the bathroom anymore. EF: Definite problem: Two respondents were not completely clear about the definition of toileting. Suggest adding a definition prior to the question such as, "Help toileting" is defined as helping in the bathroom, to get on and off the toilet, clean up after him or her, or change a resident's disposable briefs or pads. DRAFT: Revise to read, "In the last 6 months, did you help your family member with toileting in the nursing home? "Help toileting" is defined as helping to get on and off the toilet or helping to change his or her disposable briefs or pads." [Roger's suggested revision to second sentence, "Help toileting" means helping…."] See revised item in questionnaire. 4) Q26 : (nurses/aides discourage questions) had 2 respondents who (correctly) focused on the 6 month time frame but said "no" because they said nurses had "trained" them not to ask questions prior to 6 months so they didn't ask anything in the last 6 months. EF: No problem found. No changes to item. 5) Q50-52: (rating the nursing home) - this section of 3 rating DID NOT WORK for respondents - they had no idea how to distinguish the questions and felt they all were about the same thing - at Q50 that asked a rating of care - respondents included staff, medical care, physical environment, pretty much everything. Q51 seemed to be a rating of the medical care rather than a rating of confidence in medical care. At Q52 several respondents asked why we were asking the same question again! So, basically respondents felt that Q50 & Q52 were the same question and that without knowing that medical care should be considered separately, everyone included that in the rating of care (in Q50). Testing Decision:
Q50: [KEEP] Using any number from 0 to 10 where 0 is the worst care possible and 10 is the best care possible, what number would you use to rate the care at the nursing home?
Q51: [DELETE 51] Using any number from 0 to 10 where 0 is not confident and 10 is completely confident, how confident are you that your family member is receiving high quality medical care from the nursing home?
Q52. [DELETE 52] Using any number from 0 to 10, where 0 is the worst nursing home possible and 10 is the best nursing home possible, what number would you use to rate this nursing home?
(For more detail see Appendix C for First Round Cognitive Testing memo and Appendix E for Second Round Cognitive Testing memo of AIR Final report)

Summary of composite decisions: The final 4 domains are a balance of theory (original intent of items and composites), statistical evidence of reliability and validity (item-level and nursing home level, factor analyses) and stakeholder perspectives. The team, in consultation with the Technical Expert Panel eliminated ten items from the survey, two of which were negative items. In order to determine which items to keep and which items to eliminate, we closely reviewed the psychometric properties of the composites and items. We used an inter-unit reliability statistic to determine how well items and composites were able to detect differences – or discriminate – across nursing homes. We assessed the convergent validity of the items within each composite by examining the item-to-total correlations and the factor loadings. In every case these ten items had either poor convergent or discriminant validity with their composite, did not discriminate among nursing homes, and/or contained content that was included elsewhere in the survey. Two of the negative items ("Nurses and Aides Discourage Questions" and "Medical Belongings Lost" that were retained among the 21 had marginal measurement properties but were very important to the consumer

advocates and nursing home resident ombudsman in our stakeholder panel and were also indicated as a significant problem by Ombudsman data. These were included to maintain the content validity of the survey. The eight negatively-framed items were kept either because they had good or excellent measurement properties or because they discussed aspects of care that ombudsman had identified as particularly important to patients (see Table 2 in AIR Final Report on page 111).	
3b/3c. Relation to other NQF-endorsed measures 3b.1 **NQF # and Title of similar or related measures:** There are similar CAHPS measures but for different types or settings of care (Hospital CAHPS, Clinician and Group CAHPS, Home Health CAHPS). However, all the other CAHPS measures are specifically for patients, not for family members.	
(for NQF staff use) **Notes on similar/related** endorsed **or submitted measures:**	
3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source or different topic but same target population): 3b.2 **Are the measure specifications harmonized? If not, why?** Yes, the measure specifications of this CAHPS family member instrument is harmonized with other CAHPS survey measure specifications.	**3b** C☐ P☐ M☐ N☐ NA☐
3c. Distinctive or Additive Value 3c.1 **Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:** This adds a nursing home experience measure based on a family member (of a long stay resident)perspective. 5.1 **If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality:** There is no similar measure for the same target population.	**3c** C☐ P☐ M☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability*?	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, met? **Rationale:**	**3** C☐ P☐ M☐ N☐
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	Eval Rating
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 **How are the data elements that are needed to compute measure scores generated?** Survey	**4a** C☐ P☐ M☐ N☐
4b. Electronic Sources 4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*) No 4b.2 **If not, specify the near-term path to achieve electronic capture by most providers.** this is a survey instrument so electronic capture is not considered	**4b** C☐ P☐ M☐ N☐

Comment [KP23]: 3b. The measure specifications are harmonized with other measures, and are applicable to multiple levels and settings.

Comment [k24]: 16 Measure harmonization refers to the standardization of specifications for similar measures on the same topic (e.g., *influenza immunization* of patients in hospitals or nursing homes), or related measures for the same target population (e.g., eye exam and HbA1c for *patients with diabetes*), or definitions applicable to many measures (e.g., age designation for children) so that they are uniform or compatible, unless differences are dictated by the evidence. The dimensions of harmonization can include numerator, denominator, exclusions, and data source and collection instructions. The extent of harmonization depends on the relationship of the measures, the evidence for the specific measure focus, and differences in data sources.

Comment [KP25]: 3c. Review of existing endorsed measures and measure sets demonstrates that the measure provides a distinctive or additive value to existing NQF-endorsed measures (e.g., provides a more complete picture of quality for a particular condition or aspect of healthcare, is a more valid or efficient way to measure).

Comment [KP26]: 4a. For clinical measures, required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery. (e.g., BP recorded in the electronic record, not abstracted from the record later by other personnel; patient self-assessment tools, e.g., depression scale; lab values, meds, etc.)

Comment [KP27]: 4b. The required data elements are available in electronic sources. If the required data are not in existing electronic sources, a credible, near-term path to electronic collection by most providers is specified and clinical data elements are specified for transition to the electronic health record.

4c. Exclusions

4c.1 **Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?**
No

4c.2 **If yes, provide justification.**

4c
C☐
P☐
M☐
N☐
NA☐

Comment [KP28]: 4c. Exclusions should not require additional data sources beyond what is required for scoring the measure (e.g., numerator and denominator) unless justified as supporting measure validity.

4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences

4d.1 **Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.**
There could be issues if the entity collecting the data does not follow the guidelines for survey administration (e.g.,drawing the sample and assuring confidentiality). Unless the sponsor permits direct access to the resident records for random sampling, it is possible that the nursing home may select family members likely to give more favorable responses (or exclude those likely to give unfavorable responses) when selecting records for the sample. In addition, errors could be introduced if an entity adds non-Nursing Home CAHPS items before any of the core survey questions in the Nursing Home CAHPS Family Member Survey. The core survey items are all those questions prior to the "About You" section of the survey.
AHRQ has a CAHPS User Group support contract that is available to provide technical assistance for entities wishing to implement this survey- this can help reduce errors.

4d
C☐
P☐
M☐
N☐

Comment [KP29]: 4d. Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified.

4e. Data Collection Strategy/Implementation

4e.1 **Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:**
Based on the field test results that achieved more than 50% response rate in the two rounds of family surveys, our survey administration guidelines permit the phone interview phone follow up to be an optional part of administration protocol.
Additional lessons learned about obtaining the sampling frame:
After we contacted each nursing home by phone, we emailed a document describing each of the data elements and how we would like the file laid out and provided options in terms of how to provide the data. We requested two types of data: critical data elements such as the responsible party, address, phone number, and start date of care. Additionally, we provided detailed instructions about how to maintain HIPAA compliance including the need for encryption. We also followed up with calls and/or emails to see if they had questions or comments.
Despite the fact that the nursing homes were provided with detailed instructions regarding data file construction and delivery, there were several challenges to obtaining the sampling frame data. These challenges included:
- Misunderstanding on the part of the nursing homes of HIPAA rules and regulations
- Lack of technical ability on the part of the nursing homes with regard to extracting the necessary data from their systems and providing it in a usable, electronic format.

When we interacted with the nursing home, the staff we typically were in contact with were not staff who managed the data regularly. Thus, in some cases, it wasn't until later that we found out that nursing homes were having difficulty understanding which data to include. For example, we needed to know the date care began in order to determine eligibility. In some cases, some recipients had more than one start date of care, and in other cases it was unclear to an nursing home what was meant by 'start' date.
Nursing homes faxed information, emailed information, fedexed hard copies, or sent CD-ROMs. In a few cases, they had a good understanding of HIPAA polices and how to provide protected health information data safely and securely. In the instructions provided to the nursing homes, we stated that the files must be encrypted and that they could not be emailed without encrypting them first because of HIPAA rules. Nevertheless, many nursing homes emailed files or did not encrypt them, even when they thought that they had or said that they did. One nursing home was unable even to burn a file onto a CD. Some had no idea that they needed to encrypt files or how to do so. Most of the sample files were sent in a hard copy format. We used optical character recognition software to read the information (with significant quality assurance review) and/or manually keyed information. In some cases, while the information was electronic, the data

4e
C☐
P☐
M☐
N☐

Comment [KP30]: 4e. Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, etc.) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).

were not in any order or set fields, thus staff had to key in the data.
Many of the nursing homes had more than one responsible party, or one responsible party was responsible for more than one person at the nursing home. In a few cases, someone at the nursing home was the responsible party.
Lessons Learned
Below is a list of recommendations for future users:
Use Data Use Agreements: Give each nursing home 5 business days to fax a signed data use agreement. Call each nursing home 2 days before to confirm that you will get the data use agreement on the day you requested.
Confirm the data: At the same time you are obtaining the Data Use Agreements, ask them to provide an example of the data they will provide for the field test, but blacking out any personal health information or personally identifiable information. This will help you to understand whether the home is able to produce an electronic file or not and help the home prepare for developing a sample frame.
Provide information about HIPAA: Since HIPAA non-compliance was a common problem, and because both managers and IT staff need to understand HIPAA regulations, provide a short description of HIPAA including links (no more than one page). Provide links to downloading inexpensive encryption software that meets HIPAA requirements.
Contact multiple nursing home representatives. Rather than having one main nursing home contact, always ask to contact the main nursing home representative (typically an administrator, manager, or QI director) and the IT lead on the project starting from the very beginning of the project. This is helpful for two reasons; we often found that the main staff was out of the office at critical points; secondly, the IT lead is the only person who will know how the data is laid out and will understand what challenges there are to download the data. Unfortunately, this method will not work in all cases as many nursing homes do not have any IT staff.
Work with headquarters when working with chains. If you are working with an organization that is part of a chain, contact the headquarters. In many cases, the headquarter staff can facilitate the process immeasurably and even provide all of the data for each nursing home.

4e.2 **Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:**
The CAHPS Nursing Home Survey: Family Member Instrument and all composite measures are in the public domain and free to use. The costs associated with implementing these measures are the cost of the data collection, analysis and facility feedback or public reporting. The data collection budget (via subcontract with AIR) was $46,620 for survey implementation of the pilot test in 15 nursing homes (October 2006 through February 2007) for 885 completed family member surveys or $52.68 per completed survey. This cost included 2 mail rounds and phone interview followup and achieved a response rate of 66%. This does not include analysis or facility feedback costs which were borne by the CAHPS team. A similar family member survey conducted by the Maryland Health Care Commission in fall 2009 that included 2 mail rounds with postcard and a final phone reminder but no phone interviews cost $25 per completed interview and achieved a response rate of 57.7 % AHRQ's preliminary survey administration guidance for this Family Instrument is to include 2 rounds of mailing with an optional phone interview so the implementation costs could be closer to MHCC or $25 per completed interview.

4e.3 **Evidence for costs:**
Cost data are available from the AHRQ 2007 Pilot test and the 2009 MHCC Family Member Survey (similar to CAHPS). Cost data per completed survey can also be obtained from the 2008 Ohio Department of Aging Family survey and the Health Quality Council of Alberta for its 2007 family member survey (CAHPS survey).

4e.4 **Business case documentation:** The intent of the NHCAHPS initiative (also known as Nursing Home CAHPS) is to provide a standardized survey instrument and data collection methodology for measuring nursing home residents' and their family members' perspectives on nursing home care. While many nursing homes may currently collect information on patient satisfaction, prior to NHCAHPS there has been no national standard for collecting or publicly reporting nursing home residents' and family members' perspectives of care information that would enable valid comparisons to be made across all nursing homes.

In order to make "apples to apples" comparisons to support consumer choice, AHRQ has recognized the importance of creating a standard measurement approach. NHCAHPS is a core set of questions that can be combined with a broader, customized set of nursing home-specific items. NHCAHPS survey items

complement the data a nursing home may currently collect to support improvements in internal customer services and quality related activities. Three broad goals have shaped the NHCAHPS survey. First, the survey is designed to produce comparable data on the nursing home residents' and family members' perspective on care that allows objective and meaningful comparisons between nursing homes on domains that are important to them. Second, public reporting of the survey results is designed to create incentives for nursing home to improve their quality of care. Third, public reporting will serve to enhance public accountability in health care by increasing the transparency of the quality of nursing home care provided in return for the public investment. Because the government (federal and state combined) pays for almost two-thirds of the $131 billion of total nursing home costs (2008 statistics), the Centers for Medicare & Medicaid Services (CMS) are interested in the consumers' perspective on the quality of care they receive. As the federal agency responsible for nursing home quality oversight, CMS has supported the development of a consumer experience survey for both residents and their family members. With these goals in mind, the NHCAHPS CAHPS project has taken substantial steps to assure that the survey is credible, useful, and practical. This methodology and the information it generates is available to the public.	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility*?	**4**
Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? **Rationale:**	**4** **C☐** **P☐** **M☐** **N☐**
RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	**Time-limited ☐**
Steering Committee: Do you recommend for endorsement? **Comments:**	**Y☐** **N☐** **A☐**

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner)
Co.1 Organization
Agency for Healthcare Research and Quality (AHRQ/DHHS), 540 Gaither Road, Rockville, Maryland, 20850

Co.2 Point of Contact
Judith, Sangl, Sc.D., jsangl@ahrq.gov, 301-427-1308-

Measure Developer If different from Measure Steward
Co.3 Organization
Agency for Healthcare Research and Quality (AHRQ/DHHS), 540 Gaither Road, Rockville, Maryland, 20850

Co.4 Point of Contact
Judith, Sangl, Sc.D., jsangl@ahrq.gov, 301-427-1308-

Co.5 Submitter If different from Measure Steward POC
Judith, Sangl, Sc.D., jsangl@ahrq.gov, 301-427-1308-, Agency for Healthcare Research and Quality (AHRQ/DHHS)

Co.6 Additional organizations that sponsored/participated in measure development

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
AHRQ's primary way of obtaining stakeholder input was by establishing a Technical Expert Panel (TEP) composed of

industry, regulators and quality improvement organizations, payers, long- term care researchers, and consumer advocates. In addition to the Centers for Medicare & Medicaid Services, the TEP included representatives from the following organizations: (1) AARP; (2) American Health Care Association; (3) American Association of Homes and Services ; (4) National Alliance for Caregivers/Gerontology Program of Towson University ; (5) Quality Partners of Rhode Island, the CMS QIOSC for nursing home care; (6) Veterans Administration; (7) National Citizen's Coalition for Nursing Home Reform; (8) Scripps Gerontology Center of Miami University; (9) Alzheimer's Association; (10) American Medical Directors Association; and (11) National Network of Career Nursing Assistants. This TEP met in-person in November 2005 and February 2008. Additionally, AHRQ consulted with various TEP members at additional times during the survey development. Members of the TEP offered different perspectives on the use and value of a questionnaire for family members and provided valuable feedback on the proposed items, analyses, and final instrument.
Issues raised by TEP: On the February 2008 conference call, TEP members raised several issues in their review of the Final Report and survey. One issue raised was that the survey instrument does not incorporate the full range of domains of interest to consumers or facilities (dining, activities, and the admissions process were given as examples). A related point was that the survey would be difficult to use for quality improvement because it is not comprehensive. The CAHPS team responded that it is not possible to create an instrument that would serve all purposes but that the proposed questions would be useful for identifying issues that require more detailed study at the facility level. It was pointed out that it is possible to add supplemental questions to a CAHPS survey at the end of the instrument, right before the demographic items. It was further noted that the instrument was intended for family members, not to be a proxy for the resident. This family member survey should complement but not substitute for a survey of residents. Some of the topic areas considered missing in the family survey are covered in the resident survey.

Consumer advocates affirmed the importance of staff availability and staff attitude as key survey items in the recommended composites. They noted that several items proposed for deletion (e.g., "staff treating resident roughly") were frequently mentioned by family members. However, the CAHPS team explained that the items were proposed for deletion because of their poor psychometric performance. The advocates suggested that there be an explanation to family members about confidentiality because of fear of retaliation and an explanation about how the data will be used. Another suggestion was to include an open-ended question asking for additional comments on the care in the nursing home. A long term care expert recommended that the protocol materials explain how to do the case mix adjustment.

Ad.2 **If adapted, provide name of original measure:**
Ad.3-5 **If adapted, provide original specifications URL or attachment**

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.6 **Year the measure was first released:** 2008
Ad.7 **Month and Year of most recent revision:**
Ad.8 **What is your frequency for review/update of this measure?** AHRQ will work with NQF to update the measure as needed.
Ad.9 **When is the next scheduled review/update for this measure?**

Ad.10 **Copyright statement/disclaimers:** CAHPS® is a registered trademark of the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. This CAHPS® questionnaire should be used without modification to the core set of questions. Supplemental questions may be added after the core set of questions and before the demographic question section. Please consult Guidelines for Modifying and Naming CAHPS Surveys at https://www.cahps.ahrq.gov/content/products/PROD_ModifySurveys.asp

Ad.11 -13 **Additional Information web page URL or attachment:** Attachment ATTACHMENT 1 CASE MIX CALCULATION.doc

Date of Submission (*MM/DD/YY*)**:** 07/13/2010

Page 4: [1] Comment [k4]	Karen Pace	10/5/2009 8:59:00 AM

1c. The measure focus is:

- an outcome (e.g., morbidity, mortality, function, health-related quality of life) that is relevant to, or associated with, a national health goal/priority, the condition, population, and/or care being addressed;

OR

- if an intermediate outcome, process, structure, etc., there is **evidence** that supports the specific measure focus as follows:
  - Intermediate outcome – evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
  - Process – evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and
    if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
  - Structure – evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
  - Patient experience – evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
  - Access – evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
  - Efficiency – demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

ATTACHMENT 1

MEASURE SPECIFICATION SECTION
2A.21 CALCULATION ALGORITHM

PATIENT-MIX ADJUSTMENTS FOR FAMILY MEMBER NURSING HOME SURVEY

Adjustments

Specification 4 provides the steps for producing raw nursing home provider scores. Final scores shall include a case-mix adjustment to better ensure the comparability of scores across providers.

- Case-mix adjustments

The purpose of adjusting for case-mix is to estimate how different providers would be rated if they all provided care to comparable groups of patients.

- The following variables will be used in the case-mix adjustment model:
  - Respondent Age (specified as a categorical variable)
  - Respondent education (specified as a categorical variable)
  - Whether respondent perceived resident was permanently in nursing home (specified as categorical variable)
  - Whether the respondent judged the resident was capable of making decisions (specified as categorical variable)
- The case-mix adjustment uses a regression methodology also referred to as covariance adjustment. As an example:

Let y_{ipj} represent the response to item *i* of family member (respondent) *j* for resident in nursing home *p* (after recoding, if any, has been performed). The model for adjustment of a single item *i* is of the form:

$$y_{ipj} = \beta_i' x_{ipj} + \mu_{ip} + \varepsilon_{ipj}$$

where β_i is a regression coefficient vector, x_{ipj} is a covariate vector consisting of four adjuster covariates (as described above), μ_{ip} is an intercept parameter for nursing home *p*, and ε_{ipj} is the error term. The estimates are given by the following equation:

$$\left(\hat{\beta}_i' \; \hat{\mu}_i'\right)' = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_i$$

where $\mu_i = (\mu_{i1}, \mu_{i2}, \ldots \mu_{ip})'$ is the vector of intercepts, $\mathbf{y}_i$ is the vector of responses, and the covariate matrix is:

$$\mathbf{X} = \left(\mathbf{X}_a \quad u_1 \; u_2 \; \ldots \; u_p\right)$$

where the columns of $\mathbf{X}_a$ are the vectors of values of each of the adjuster covariates, and u_p is a vector of indicators from nursing home p, p = 1, 2,…P,

with entries equal to 1 for family members (respondents) of residents in nursing home p and 0 for others.

The estimated intercepts are shifted by a constant amount to force their mean to equal the mean of the unadjusted nursing home means $\bar{y}_{ip}$ (to make it easier to compare adjusted and unadjusted means), giving adjusted nursing home means:

$$\hat{a}_{ip} = \hat{\mu}_{ip} + (1/P)\sum_p \bar{y}_{ip} - (1/P)\sum_p \hat{\mu}_{ip}$$

For single-item responses, these adjusted means are reported. For composites, the several adjusted means are combined using the mean:

$$\hat{a}_p = (1/P) \Sigma_i \hat{a}_{ip}$$

CODEBOOK CAHPS NURSING HOME FAMILY MEMBER SURVEY SPECIFICATIONS

Composite 1. Meeting Basic Needs – Help with Eating, Drinking and Toileting

In this composite, all three questions use the "yes" or no" response, in which "no" represents better quality.

Q18: Family member helped nursing home resident with eating. "Was it because the nurses or aides either didn't help or made him or her wait too long?"

- Yes=1
- No=2

Q21: Family member helped nursing home resident with drinking. "Was it because the nurses or aides either didn't help or made him or her wait too long?"

- Yes=1
- No=2

Q23: Family member helped nursing home resident with toileting. "Was it because the nurses or aides either didn't help or made him or her wait too long?"

- Yes=1
- No=2

COMPOSITE 2. Nurses and Aides' Kindness and Respect towards Resident

In this composite, four questions use the "never," "sometimes," "usually," and "always" response scale, and one question uses a "yes" or "no" rating, in which a higher number represents better quality.

Q12: "In the last 6 months, how often did you see the nurses and aides treat your family member with courtesy and respect?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

Q13: "In the last 6 months, how often did you see the nurses and aides treat your family member with kindness?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

Q14: "In the last 6 months, how often did you feel that the nurses and aides really cared about your family member?"

- Never=1
- Sometimes=2

- Usually=3
- Always=4

Q15: "In the last 6 months, did you ever see any nurses or aides be rude to your family member or any other resident?"

- Yes=1
- No=2

Q26: "In the last 6 months, how often did the nurses and aides handle the situation in a way that you felt was appropriate?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

COMPOSITE 3. How Well the Nursing Home Provides Information and Encourages Family Involvement

In this composite, four questions use the "never," "sometimes," "usually," and "always" response scale, and two questions use a "yes" or "no" response, in which a higher number represents better quality.

Q24: "In the last 6 months, how often did you get this information as soon as you wanted?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

Q30: "In the last 6 months, how often did the nurses and aides explain things in a way that was easy for you to understand?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

Q31: "In the last 6 months, did the nurses and aides ever try to discourage you from asking questions about your family member?"

- Yes=1
- No=2

Q43: "In the last 6 months, did you ever stop yourself from talking to any nursing home staff about your concerns because you thought they would take it out on your family member?"

- Yes=1

- No=2

Q45: "In the last 6 months, how often were you involved as much as you wanted to be in the decisions about your family member's care?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

Q53: "In the last 6 months, how often did you get all the information you wanted from the nursing home about payments or expenses?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

COMPOSITE 4. Nursing Home Staffing, Care of Belongings, and Cleanliness

In this composite, four questions use the "never," "sometimes," "usually," and "always" response scale, and two questions use the "never," "once," or "two or more times" ratings, in which a higher number represents better quality.

Q.11: "In the last 6 months, how often were you able to find a nurse or aide when you wanted one?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

Q24: "In the last 6 months, how often did your family member's room look and smell clean?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

Q32: "In the last 6 months, how often did your family member look and smell clean?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

Q37: "Personal medical belongings are things like hearing aids, glasses, and dentures. In the last 6 months, how often were your family member's personal medical belongings damaged or lost?"

- Two or more times=1
- Once=2
- Never=3

Q39: "In the last 6 months, when your family member used the laundry service, how often were clothes damaged or lost?"

- Two or more times=1
- Once=2
- Never=3

Q51: "In the last 6 months, how often did you feel there were enough nurses and aides in this nursing home?"

- Never=1
- Sometimes=2
- Usually=3
- Always=4

Global Rating and Items

One question uses the 0 to 10 rating scale in which "10" represents the highest quality and "0" represents the lowest quality; one question uses a "yes," "no" scale in which a higher number represents better quality, and one question uses a "definitely no," "probably no," "probably yes," and "definitely yes" scale in which a higher number represents better quality.

Q40: "In the last 6 months, were you ever unhappy with the care your family member received at the nursing home?"

- Yes=1
- No=2

Q49: "What number would you use to rate the care at this nursing home?"

- 10 = 10 (highest quality)
- 0 = 0 (lowest quality)

Q50: "If someone needed nursing home care, would you recommend this nursing home to them?"

- Definitely No=1
- Probably No=2
- Probably Yes=3
- Definitely Yes=4

AMERICAN INSTITUTES FOR RESEARCH®

CAHPS® Survey for Family Members of Nursing Home Residents

Final Report

July 2008

Prepared for:

Judith Sangl, ScD.
Project Officer
Agency for Healthcare Research and Quality
540 Gaither Road, 3rd Floor
Rockville, MD 20850

Prepared by:

Elizabeth Frentzel, M.P.H.
Chris Evensen, M.S.
San Keller, Ph.D.
Steven A. Garfinkel, Ph.D.
American Institutes for Research
101 Conner Drive, Suite 301
Chapel Hill, NC 27514

Executive Summary

The CAHPS II team, comprising the American Institutes of Research (AIR), the RAND Corporation, and Yale University/Harvard University/University of Massachusetts Center for Survey Research (CSR), with assistance from Westat, developed, tested, and fielded a nursing home survey for family members or responsible parties. The CAHPS Nursing Home *Family* Survey ("family survey") complements the CAHPS Nursing Home *Resident* Survey ("resident survey") in that it was not designed to provide the same information that the resident survey provides. The family survey does not ask respondents to report on the residents experience; rather, it asks respondents to report on their own experiences.

Survey Development Process

The development of this survey instrument involved several steps:

1. Publishing a call for measures in the Federal Register, including requesting survey, items or domains from other home health surveys
2. Reviewing previous literature and other nursing home surveys
3. Conducting and focus groups with family members of nursing home residents.
4. Developing a draft survey instrument
5. Cognitively testing the draft survey in English
6. Obtaining input from the nursing home technical expert panel (TEP) to review candidate items and composites
7. Refining the survey instrument
8. Pilot testing the survey with 15 nursing homes in Texas
9. Analyzing the pilot data and revising the survey
10. Obtaining input from the TEP to finalize the survey.

The composite structure was determined using psychometric statistical techniques, including factor analysis and multi-trait analysis. The case-mix adjustment analysis included multiple regression and variance component analyses.

Final Composite Structure

As part of the standard CAHPS survey development process, composite measures are developed from items measuring the same aspect of care. The CAHPS team conceptually defined

composites representing our hypothesis about the "structure" that the survey data would reflect. We then conducted a confirmatory factor analysis (CFA) to determine whether the pilot data were consistent with the hypothesized structure. The CFA did not wholly support the structure we had hypothesized. Therefore, we conducted an exploratory factor analysis to help us define reliable and valid composites and to help us identify items that should be revised or deleted. The composite structure that AIR proposed and the team agreed to presents the best combination of conceptual properties and statistical support. The proposed composite structure is listed in Table 1 (this is also in Table 7 on page 13 in full report).

Table 1. Final Composites and Items

Q#	Composite or Item Handle
	Meeting Basic Needs: Help with eating, drinking, toileting [a]
19	Wait too long for help with eating
21	Wait too long for help with drinking
23	Wait too long for help with toileting
	Nurses/Aides' Kindness/ Respect Towards Resident
12	Nurses/Aides treat resident with respect
13	Nurses/Aides treat resident with kindness
14	Nurses/Aides really cared about. resident
15	Nurses/Aides rude to resident
26	Nurses/Aides appropriate with violent resident
	Nursing Home Provides Info/ Encourages Respondent Involvement
29	Nurses/Aides give respondent information about resident
30	Nurses/Aides explain things to respondent
31	Nurses/Aides discourage respondents questions
43	Respondent stops self from complaining
45	Respondent involved in decisions about care
53	Respondent given Info. about payments/expenses
	Nursing Home Staffing, Care of Belongings, & Cleanliness
11	Can find a nurse or aide
51	Enough nurses/ aides?
32	Room looks/smells clean
24	Resident looks/ smells clean
35	Public areas look/smell clean
37	FM's medical belongings lost
39	FM's clothes lost

This composite structure excludes 10 substantive items that did not perform well. Table 2 lists the excluded items and explains why they were excluded.

Table 2. Excluded Items

Q#	Items Excluded from Composites	Reasons for Excluding
16	Nurses/Aides treat resident rough	Marginal measurement characteristics.
17	Another resident rude to respondent's family member	Marginal measurement characteristics.
27	Nurses/Aides treat respondent with respect	Not uniquely and strongly related to its composite. Other items tap concept.
33	Noise level around room	Not uniquely and strongly related to its composite.
34	Places to talk to resident in private	Not uniquely and strongly related to its composite. Other items tap concept.
36	Nurses/Aides didn't protect resident's modesty	Marginal measurement characteristics.
42	Staff handled respondent's concerns satisfactorily	Includes term "satisfaction" which is not consistent with CAHPS methods; also low NH-reliability
46	Nursing Home has care conferences	Care conferences are required in Canada ONLY.
48	Management handled complaints satisfactorily	Includes term, "satisfaction" which is not consistent with CAHPS methods; also low NH-reliability
58	Staff did not give respondent info. due to privacy laws	Marginal measurement properties.

Recommended Case-Mix Adjusters

After conducting numerous analyses, we concluded that using four variables should be used as case-mix adjusters for the family survey: respondent age, respondent education, whether resident was permanently in the nursing home, and whether the resident was capable of making decisions.

Contents

The Development of a CAHPS® Survey for Family Members of Nursing Home Residents

Project Description

Over the last decade, the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) program, has led the way in developing a set of valid and reliable surveys that ask consumers and patients to report on and evaluate their experiences with health care. The findings from these surveys have been used to inform consumer healthcare choices and for quality improvement activities at the facility level. The CAHPS Nursing Home Family survey is an addition to this family of products. Research and development for this survey was sponsored by the Centers for Medicare and Medicaid Services and the Agency for Healthcare Research and Quality. The objective of the survey is to collect data that measure family members' perceptions of nursing home care.

The CAHPS Nursing Home *Family* Survey ("family survey") complements the CAHPS Nursing Home *Resident* Survey ("resident survey") in that it was not designed to provide the same information that the resident survey provides. The family survey does not ask respondents to report on the residents experience; rather, it asks respondents to report on their own experiences. In other words, the family survey does not ask the respondent to provide proxy responses for the resident's experience. This is an important distinction because research has demonstrated that queries of family members do not provide the same data as queries of residents on nursing home satisfaction and experience-of-care surveys.

Description of Tasks and Report Structure

This report describes the ten step process to develop the survey grouped into the following four activities:

1. **Formative research.** Formative research activities included three activities; publishing a call for measures in May 6, 2004, a literature review, and focus groups with family members of nursing home residents.

2. **Development of items and cognitive testing**. The next two activities included drafting an initial set of survey items and two round of cognitive testing with family with family members of nursing home residents.

3. **Technical Expert Panel and refinement of instrument.** The next two steps were to hold a Technical Expert Panel (TEP) and refine the survey as a result. The TEP met on November 2, 2005 to review candidate items and composites. The TEP included representatives from the nursing home industry, regulators and quality improvement organizations, consumers, providers, and long term care researchers.

4. **Pilot survey, analysis, TEP review, and finalization of instrument**. The instrument was field tested in east Texas. The Texas State Long Term Care Ombudsman recruited 15 Texas nursing homes and the Health Quality Council. The CAHPS Consortium conducted the Texas field test. In addition, the Health Quality Council of Alberta conducted an independent field test in Alberta and contributed their data to AIR for the psychometric analysis conducted by AIR. The TEP provided a review after wards to ensure that necessary substantive items were included.

In the next sections, we describe each of these four activities.

I. Formative Research

The formative research included focus groups, a call for measures, and reviewing literature.

Focus Groups

The goals of the focus groups were to:

- Understand participants' current experiences with nursing homes.
- Determine how participants' conceptualize good care.
- Determine the comparative salience and importance of the factors associated with good care.
- Understand participants' potential uses of a nursing home quality report.

In order to answer these questions, the three CAHPS grantees (Harvard University, AIR, and RAND, conducted 12 focus groups: two each in New York City and Phoenix, Arizona and four each in Palo Alto, California and Chapel Hill, North Carolina. With the exception of those in North Carolina, all focus group participants had already chosen a nursing home for a family member. Those participants in North Carolina were at the stage of considering moving a relative to a nursing home within a year. Findings from these focus groups can be found in ***Appendix A, Focus Group Findings from AIR, RAND, and Harvard.***

The call for measures and the literature review were conducted as part of the initial Nursing Home *Resident* Survey and are not included here.

II. Development of Items and Cognitive Testing

Item Development

We used the results from the call for measures, the literature review, and the focus groups, to develop a list of initial domains and questionnaire items within each domain (see ***Table 3. Draft Domains and Item Handles Prior to Cognitive Testing***).

Table 3. Draft Domains and Item Handles Prior to Cognitive Testing

Domains and Item Handles
1. Domain: Communication
• Converse with staff about family member's care
• Courtesy and respect
• Language differences
• Obtaining information about family member's medical condition
• Nurses/aides speak to you about how resident doing/feeling
• Comfort with asking questions about family member
• Getting information about payments and expenses
• Management availability
• Staff ability to manage concerns if there were concerns
• Concern about staff taking out your issues on family member
• Stop yourself from talking to staff because of concern about staff taking it out on resident
2. Domain: About Staff Behavior (psychosocial)
• Really care about resident
• Treat resident with courtesy
• Treat resident with kindness
• Staff checking on residents
• Staff rude to residents
• Staff treat residents roughly
• Family member ever complained to you about being physically abused or treated roughly by staff?
• Family member ever complained to you about being physically abused or treated roughly by other residents?
• Waiting for eating because of staff delay
• Waiting for drinking because of staff delay
• Waiting for dressing, toileting, or bathing because of staff delay
3. Domain: About the Nursing Home (Environment)
• Problems with laundry service (if applicable)

•	Cleanliness of room
•	Working order of equipment
•	Acceptability of noise level near room
•	Ability to find places to talk in private
•	Cleanliness of public areas
•	Unpleasant smells
4.	**Domain: Care of Your Family Member**
•	Family member look and smell clean
•	Staff help family member when needed
•	Effort to be gentle by aides and nurses
•	Nurses and aides tell family member what to expect
•	Protect modesty
•	Encouraging of family member to participate in care decisions
•	Encourage family member to be independent
•	Same nurses and aides provide care to family member
•	Enough involvement in care
•	Staff consult you about family member's care
•	Participation in care conference
•	Concern about food intake
•	Concern about hydration
5.	**Ratings**
•	Keeping track of family member's belongings
•	Ability to make room home like
•	Public areas homelike
•	Availability of nurses and aides
•	Care given by nurses and aides
•	Care given by nurse practitioners and doctors
•	Management
•	Overall nursing home rating
•	Recommendation

Cognitive Testing

We prepared a formatted "cognitive testing" version of the instrument and a cognitive interviewing protocol for use by the interviewers. This protocol provided a listing of scripted probes that could be employed to provide insights into each respondent's cognitive processes as he or she read and answered the pilot items. It also included a series of general questions

about the items, to allow the respondent to provide additional feedback about the items and to help assess the comprehensiveness of the instrument. A think aloud training exercise, with practice questions and a scripted response for the interviewers to use in modeling appropriate thinking aloud behaviors, was also included (see ***Appendix B: Draft Survey and First Cognitive Testing Protocol***).

The cognitive interviews examined the following issues related to the draft questionnaires:

1. **Content**: Are the questions that are included in the survey important to consumers? Are consumers able to make judgments about the questions?

2. **Comprehension and Interpretation**: Are the words, phrases, and questions easy to understand regardless of education level or knowledge of nursing home care? Are the questions interpreted as intended?

3. **Recall:** Are consumers able to recall the events asked about and to make judgments about them?

4. **Navigation:** Did the survey flow correctly? Were people skipped out of sections appropriately?

The results of the first round of cognitive testing are included in ***Appendix C: First Round Cognitive Testing Memos.*** Following the first round of cognitive tests, we revised the items and prepared a protocol for the second round of cognitive testing (see ***Appendix D: Draft Survey and Second Cognitive Testing Protocol***). This round tested the items as both self-administered items and as interviewer-administered items under the assumption that the final instrument would probably be administered by both mail and telephone. The results of the second round of testing are included in ***Appendix E: Second Round Cognitive Testing Memos.*** The instrument was revised again after the second round of cognitive testing.

Anyone with a family member in a nursing home was eligible to participate in the cognitive interviews; however persons were selected so as to assure variation in race, ethnicity, and education. The team conducted a total of 27 interviews in the first testing round in June 2005 and conducted another 27 interviews in the second round in June 2006.

Trained cognitive interviewers conducted one-on-one, in-person interviews. Using a "think-aloud" approach for the interviews, we asked the participant to read each question (or read the

question verbally to the participant), provide a verbal response, and explain the reason for the response. We then followed up with probes after each of the questions to ask about specific potential problems with each item. Immediately after each interview, the interviewer wrote a summary of the participant's comments. A member of each grantee team reviewed these summaries and extracted common themes for each item. At the end of each round, the team met to review these themes and make recommendations to the larger family survey team. The larger team made decisions together regarding each of the items.

III. Technical Expert Panel and refinement of instrument

Technical Expert Panel

Between the two rounds of testing, the team met with nursing home advocacy organizations to obtain their guidance and suggestions about the draft survey. The Technical Expert Panel included representatives from the following organizations:

- AARP
- American Health Care Association
- National Network of Career Nursing Assistants
- American Association of Homes and Services
- PIOSC: Quality Partners of Rhode Island
- Veterans Administration
- National Citizen's Coalition for Nursing Home Reform
- Scripps Gerontology Center of Miami University
- Alzheimer's Association
- American Medical Directors Association
- Gerontology Program of Towson University

Refinement of instrument

Based on the initial cognitive testing findings and suggestions, we made revisions to the instrument and tested the instrument again. After the second round of cognitive testing, the instrument underwent additional revisions (see ***Appendix E: Second Cognitive Testing Memos***). The final domains and item handles are listed in ***Table 4, Draft Domains and Items Used for Pilot Study***, below. The survey instrument that we produced following both rounds of cognitive testing is included as ***Appendix F, Pilot Study Survey*).**

Table 4. Draft Domains and Items Used for Pilot Study

Domains and Item Handles
1. Domain: Getting Care Quickly
• In the last 6 months, during any of your visits, did you try to find a nurse or aide for any reason? Yes/No (**Y/N**)
• In the last 6 months, how often were you able to find a nurse or aide when you wanted one? Never, Sometimes, Usually, or Always (**NSUA**)
• In the last 6 months, during any of your visits, did you help your family member with eating? Y/N → skip
o Was it because the nurses or aides either didn't help or made him or her wait too long? Y/N
• In the last 6 months, during any of your visits, did you help your family member with drinking? Y/N→ skip
o Was it because the nurses or aides either didn't help or made him or her wait too long? Y/N
• Help toileting includes helping someone get on and off the toilet or helping change disposable briefs or pads. In the last 6 months, during any of your visits, did you help your family member with toileting? Y/N →skip
o Was it because the nurses or aides either didn't help or made him or her wait too long? Y/N
• In the last 6 months, how often did you feel there were enough nurses and aides in this nursing home? NSUA
2. Domain: Quality of Care by Nurses and Aides
• In the last 6 months, how often did you see the nurses and aides treat your family member with courtesy and respect? NSUA
• In the last 6 months, how often did you see the nurses and aides treat your family member with kindness? NSUA
• In the last 6 months, how often did you feel that the nurses and aides really cared about your family member? NSUA
• In the last 6 months, did you ever see any nurses or aides be rude to your family member or any other resident? Y/N
• In the last 6 months, did you ever see any nurses or aides treat your family member or any other resident roughly? Y/N
• In the last 6 months, how often did your family member look and smell clean? NSUA
• In the last 6 months, did you ever see the nurses and aides fail to protect any resident's privacy while the resident was dressing, showering, bathing, or in a public area? Y/N
• Sometimes residents make it hard for nurses and aides to provide care by doing things like yelling, pushing, or hitting. In the last 6 months, did you see any resident, including your family member, behave in a way that made it hard for nurses or aides to provide care? Y/N→ skip
o In the last 6 months, how often did the nurses and aides handle the situation in a way that you felt was appropriate? NSUA
3. Domain: Communication of Nurses and Aides
• In the last 6 months, how often did the nurses and aides treat you with courtesy and respect? NSUA
• In the last six months, did you want to get information about your family member

from a nurse or an aide? Y/N → skip
    - In the last 6 months, how often did you get this information as soon as you wanted? NSUA
- In the last 6 months, how often did the nurses and aides explain things in a way that was easy for you to understand? NSUA
- In the last 6 months, did the nurses and aides ever try to discourage you from asking questions about your family member? Y/N

4. Domain: Communication--Other Staff and Administrators

- In the last 6 months, were you ever unhappy with the care your family member received at the nursing home? Y/N → skip
    - In the last 6 months, did you talk to any nursing home staff about this? Y/N → skip
    - In the last 6 months, how often were you satisfied with the way the nursing home staff handled these problems? NSUA
- In the last 6 months, did you ever stop yourself from talking to any nursing home staff about your concerns because you thought they would take it out on your family member? Y/N
- In the last 6 months, have you been involved in decisions about your family member's care? Y/N→ skip
    - In the last 6 months, how often were you involved as much as you wanted to be in the decisions about your family member's care? NSUA
- In the last 6 months, did you need to contact the nursing home administrator about any problems? Y/N→ skip
    - In the last 6 months, how often were you satisfied with the way he or she handled these problems? NSUA
- A care conference is a formal meeting about care planning and health progress between a care team and a resident and his or her family. In the last 6 months, have you been part of a care conference? Y/N
- In the last 6 months, did you ask the nursing home for information about payments or expenses? Y/N → skip
    - In the last 6 months, how often did you get all the information you wanted from the nursing home about payments or expenses? NSUA
- In the last 6 months, has any nursing home staff told you that they could not give you information about your family member because of privacy laws? Y/N

5. Domain: Nursing Home Environment

- In the last 6 months, how often did your family member's room look and smell clean? NSUA
- In the last 6 months, how often was the noise level around your family member's room acceptable to you? NSUA
- In the last 6 months, how often were you able to find places to talk to your family member in private? NSUA
- In the last 6 months, how often did the public areas of the nursing home look and smell clean? NSUA
- In the last 6 months, did you ever see a resident be rude to or treat your family member or any other resident roughly? Y/N
- Personal medical belongings are things like hearing aides, glasses, and dentures. In the last 6 months, how often were your family member's personal medical belongings damaged or lost? Never/once/two or more times
- In the last 6 months, did your family member use the nursing home's laundry service

for his or her clothes? Y/N→ skip
o In the last 6 months, when your family member used the laundry service, how often were clothes damaged or lost? Never/once or twice/3+ times
6. Ratings
• Using any number from 0 to 10 where 0 is the worst care possible and 10 is the best care possible, what number would you use to rate the care at this nursing home? 0-10
• If someone needed nursing home care, would you recommend this nursing home to them? Definitely no/probably no/ probably yes/ definitely yes

IV. Pilot survey, analysis, TEP review, and finalization of instrument

Pilot Survey

The pilot survey was conducted in Texas. Health Quality Council of Alberta had its own field test and supplied data to AIR for psychometric analyses. The Canadian instrument was slightly modified as a result of differences in the organization of nursing home care in Canada and the U.S. The following description of the methods focuses on the Texas pilot study.

Sampling Frame

Nursing Home Recruitment

AIR worked with its subcontractor, Texas A&M University, to identify a diverse range of nursing homes for the pilot survey. We were assisted by the Texas Ombudsman Office within the Texas Department of Aging and Disability Services. Ombudsmen recruited 18 nursing homes from the Dallas and East Texas regions. We recruited Medicare and Medicaid certified nursing homes that represented a range of quality scores in the Texas Quality Reporting System (**QRS)**. Nursing facilities that accept Medicaid or Medicare are compared in the QRS on the basis of four dimensions of quality: technical quality, quality of life, regulatory compliance, and customer satisfaction. The mean score on the 1-100 scale is 56. Of the 20 nursing homes that initially agreed to participate, 15 nursing homes actually participated in the survey. Of the five nursing homes which opted out, reasons for declining included: 1) concern for respondent burden (they limit number of family surveys to one every 6 months); 2) the nursing home was closing down within a month; 3) low number of eligible patients; 4) change in key nursing home staff; and 5) unable to provide contact information for family members in time for survey.

Of the 15 participating nursing homes, 5 had QRS scores below 56, the average, one facility had a score of 56, and the other 9 exceeded the average. The average rating of participants was 62.6, slightly above the overall statewide mean of 56. The average rating for the lower-scoring nursing homes is 30.2 and the average rating for the higher-scoring nursing homes was 81.4 (not including the nursing home scored at 56.

Sampling Frame Eligibility

Eligible sample members were identified first by identifying residents who met the eligibility criteria and then identifying the responsible person for those residents. An eligible sample member was the person listed by the nursing home as the responsible party for a resident who had resided at the facility for at least 30 days. In addition to family members and friends, guardians, medical powers of attorney, and attorneys were considered to be eligible responsible parties. Once we received the list of eligible responsible parties from the nursing home, we excluded or re-designated the eligible responsible party when, in our judgment, the named party would have been uninformed or biased about the nursing home's performance.

1. In 31 cases, representing 2.1% of the responsible parties received from the nursing homes, the responsible party was a nursing home resident (including self) or a nursing home staff member or a trust fund or another institution. All of these were excluded.

2. In 28 cases (1.9%), the responsible party listed was the resident or the nursing home, but there was another, eligible contact listed as well. In these cases, the other contact was used instead of excluding the case.

3. In 44 cases (3.0%), residents had more than one eligible responsible party listed. This does not include the other contacts most residents had listed; only cases in which multiple persons were listed as responsible parties for the same resident. In these cases, one person was randomly selected to participate in the survey.

4. Another 17 eligible sample members (1.2%) were listed as the responsible party for more than one resident. For these cases, we randomly selected a resident for them to respond about.

5. In 12 cases (0.8%), an eligible responsible party was listed without contact information, but another person listed as a contact (not a responsible party) had adequate contact information. In these cases, we used the alternative with adequate contact information.

In addition to the above exclusions based on administrative data, the questionnaire included several screening questions that excluded other cases. It screened out eligible sample members who had visited the focal resident less than once in the last six months, and whose

focal resident had been discharged from the focal nursing home at the time of the survey. Participants with family members who were discharged were excluded, because it became apparent during cognitive testing, that these were two different audiences and a separate instrument was needed to examine experiences of people who had a nursing home resident recently discharged or transferred from a nursing home.

Sampling Methodology

Each participating nursing home forwarded a data file containing the following items for each eligible respondent: name of responsible party, address, phone number, resident/patient name, date-of-birth, gender, relationship of responsible party to resident, whether responsible party was the power of attorney, admission date, and whether the resident was in a dementia unit.

Our goal was to sample 150 potential respondents per nursing home, but the average number of eligible frame members was approximately 90 per nursing home. When the frame contained fewer than 150 eligible sample members, we sampled everyone. In one case, the number exceeded 150 by a small amount and we also sampled everyone. The initial sample size was 1,471; after removing those without an address, the sample fell to 1,444. The survey began in late October 2006 and ended early January 2007. All eligible respondents were mailed the family survey questionnaire, with a cover letter and a return postage-paid envelope. A reminder postcard was sent approximately 2 weeks after the first questionnaire was mailed. A second questionnaire and cover letter was sent approximately 2 weeks after the postcard was sent. Approximately 2 weeks after that, computer-assisted telephone interviews (CATI) began for nonrespondents. A maximum of five telephone attempts were made on different days and times of the day to try to maximize response rates. The final response rate was 66 percent (N = 885). Response rates differed by wave and mode:

- Wave 1 (initial survey): 42%
- Wave 2 (second survey): 14%
- Phone follow-up: 10%

Analysis

The AIR team, with review by RAND and Yale team members, conducted multiple activities to analyze the data for validity, reliability, and case mix. The analysis included the data from the 14 nursing homes in the Alberta pilot survey that had taken place two months earlier. We generated descriptive statistics at the item level and identified missing data, out of range values, and skip pattern inconsistencies (***Appendix G, Frequencies from pilot study)***. We also generated descriptive statistics presenting the Texas and Alberta data. The instrument team met and evaluated each item to determine whether they were similar enough to be used together analytically. In most cases, and in all cases of critical content, we found a high level of convergence at the item level.

Item Functioning

Survey questionnaires that are poorly designed – that is, those with complicated skip patterns, hard-to-understand instructions, a readability level that is beyond many respondents, or that are otherwise too complex – may yield unreliable results. We can assess how well designed a survey is by observing how well the items on the survey 'function.' The functioning is measured by analyzing patterns of inconsistent or missing responses to survey items, such as a respondents' failure to follow skip instructions or the relative amount of missing data.

The family survey contained several skip patterns; each skip consists of a gate, or screener, and one or more items controlled by that gate item. To assess how well the survey functioned, we examined every gate-item pair (i.e., skip) in the survey, including nested skips. We also examined any items that were not part of a skip pattern to check for missing data – in all, we assessed 74 items (Q02 through Q66)[1] and assigned one of five dispositions to every item, for every respondent:

1. Correct Skip (CS) – the respondent (correctly) skipped an item that they were supposed to skip.

[1] Note: the total number of items is greater than the number of questions because some items, such as race (Q63) and proxy (Q66), are 'code-all-that-apply' items, so a single item actually produces multiple variables, each of which indicates whether or not the respondent marked one of the response options.

2. Failed Skip (FS) – the respondent answered an item that they should have skipped.
3. Indeterminate Eligibility (IE) – the respondent left the gate preceding the item blank, so eligibility to answer the item cannot be determined.
4. Truly Missing (TM) – the respondent was eligible to answer the item, but left it blank.
5. Legitimate Response (LR) – an eligible respondent gave a legitimate answer to the item.

For each item, we calculated the percentage of respondents who fell into each of these five dispositions. Table 5 provides and example of the distribution of these dispositions for question 11 on the survey:

Table 5. Item Functioning Dispositions for Q11

Disposition	Frequency	Percent
Correct Skip	141	16.53
Failed Skip	32	3.75
Indeterm Elig	13	1.52
Truly Missing	8	0.94
Legit Resp	659	77.26
Total	**853**	**100.00**

We focused in particular on the percentage of respondents who had failed skips and the percentage of respondents with missing responses for each item to identify items that respondents may have had trouble completing. High rates of missing data or skip failures for an item may indicate that the question was confusing, too personal, or offensive. Our analysis showed that the rate of truly missing was less than 5% for all survey items that were intended to be used for reporting (Q04 – Q51). The rate of failed skips was less than 10% for all survey items that were part of a skip pattern.

We calculated some summary measures for the item dispositions by creating count variables that sum each occurrence of each of the five dispositions across all 853 respondents. For example, the number of items an ineligible respondent failed to skip is summed for each respondent, then that total is summed across all 853 respondents, and the mean calculated.

The same is done for the other dispositions. Out of 74 survey items, on average, respondents made fewer than one skip error, had slightly more than two truly missing responses, and had slightly more than two items where we could not determine their eligibility to respond.

We also calculated an item-level response rate (IRR), which, for each respondent, is equal to the number of items with a legitimate response divided by the number of items that the respondent was eligible to answer.

The rate of item non-response (RINR) is the complement to the IRR, or:

$$RINR = 1 - IRR$$

The mean item-level response rate for all 74 items across all respondents was 93.6%.

Table 6 gives the mean item-level response rate by nursing home and the percent of respondents at each facility who had non-missing data for at least 90% of the items they were eligible to answer (i.e., their rate of item non-response was 10% or less).

Overall, item nonresponse for this survey instrument was comparable to what was found with the pilot CAHPS Hospital survey, which was in the range of 2 to 4 percent (Elliott, Edwards, Angeles, Hambarsoomians & Hays, 2005).

Table 6. Mean Item-Level Response Rate by Facility

Nursing Home Number	Mean Facility-level IRR	% Respondents with IRR of 90% or Greater
1	91.3%	75.6%
2	94.8%	90.0%
3	93.5%	90.0%
4	89.9%	72.1%
5	96.1%	94.3%
6	92.1%	85.2%
7	96.4%	93.8%
8	92.0%	77.8%
9	93.9%	82.5%
10	94.2%	93.1%
11	93.6%	84.4%
12	94.9%	91.2%
13	92.2%	85.7%
14	93.2%	80.0%
15	91.0%	76.9%
Total	**93.6%**	**86.3%**

The combined problem of missing data and skip errors was a noticeable problem in one section of the survey that includes questions that ask about legal matters related to the residents care (questions 56, 57 and 58[2]), none of which are used in the composites. The rate of item nonresponse was 24% for Q56 and 23% for Q58, and the rate of failed skips for Q57 was over 7%.

[2] Q56: A court-appointed legal guardian has the authority to make decisions for another person because the other person cannot make decisions. Are you your family member's legal guardian? [if Yes, Go to Question 58]
Q57: Is someone else your family member's legal guardian? [Yes, No, Don't Know]
Q58: In the last 6 months, has any nursing home staff told you that they could not give you information about your family member because of privacy laws? [Yes, No]

Data Cleaning

We followed standard CAHPS data cleaning rules:

1. For gate-item pairs where the gate was a yes-no question and the item was a 'how often' question: If the gate was left blank, and the follow-up item was answered 'never', we coded the follow-up as missing. If the follow-up was answered SUA (sometimes, usually, or always), we kept the response to the follow-up, and back-coded the gate question to 'Yes'.

2. If a gate question was missing (blank/not ascertained), and subsequent survey items controlled by that gate question contained valid responses, the responses for those items were retained.

3. Failed skips: If the response to a gate question was valid, but the respondent violated the skip instruction by answering survey items that should have been skipped, the response to the gate question was retained and the responses for survey items within the gate were set to missing.

4. Correct skips were coded as missing.

Psychometric Analysis

Prior to data collection and in consultation with stakeholders, we organized 31 of the substantive survey items into a set of five domains. The first step in the psychometric analysis was to test whether there was empirical evidence to support the hypothesized item-domain relationships. We combined the Alberta and Texas data, and used that combined data set to calculate a Cronbach's alpha for each of the five domains. The alphas for 3 out of the 5 domains were greater than the standard 0.70, and the alphas for two of the domains were lower. While these results provide some support for the hypothesized relationships, the item-total correlations for the domains indicated that each domain's alpha would improve if certain items were dropped, which indicates that it was likely that there was a better item-domain

structure to be specified. Separate analyses of the Alberta and Texas data replicated these findings.

Our next step was to make use of some exploratory methods to identify composites. To make use of all available data, we obtained maximum likelihood estimates of the covariance matrix under the Missing at Random (MAR) model using a multiple imputation procedure (MI, Rubin 1976, 1987).[3] Peer-reviewed publications involving another CAHPS instruments illustrate the use of the MAR model and SAS PROC MI for this purpose (Hurtado, Angeles, Blahut & Hays, 2005; Keller et al., 2005; O'Malley et al., 2005).

We conducted an exploratory factor analysis (EFA) on the correlation matrix produced by PROC MI. The EFA used the principle factor method with squared multiple correlations as initial communality estimates, and oblique rotation (promax) with Kaiser normalization. The number of factors was determined by the eigenvalues and the interpretability of the rotated factor pattern matrix. The EFA results did not demonstrate a definitive underlying factor structure for these 31 items. The analysis proceeded at this point through an iterative process that included additional factor-analysis and multi-trait multi-item analysis, along with conversations among various members of the analysis team and CMS.

The team agreed on a final set of 21 items organized into four domains. The first three composites are more specific in focus and the fourth, more general. The first composite refers to whether respondents perceive that nurses and aides provide help with basic needs of residents who require that type of help. Only those respondents who helped their family member with eating, drinking, or toileting were eligible to answer the questions in the first composite (approximately 30% of the respondents). As might be expected, the residents that these respondents visited tended to be more impaired than the residents visited by respondents who did not help with eating, drinking and toileting (data not reported but available upon request). For convenience, we refer to this composite as "Meeting Basic Needs."

[3] The MAR model and SAS PROC MI in particular has been recommended as a method of producing a complete data file preliminary to analyses on the underlying structure of questionnaires (Tomarken & Waller, 2005). By producing multiple estimates for each missing data value, MI does not restrict the total variance in the data matrix as does a single imputation procedure.

The second composite refers to the interpersonal manner in which the nurses and aides interact with residents, "Nurse/Aide Kindness", for short. The third composite refers to how much the nursing home shares information related to the resident's care with the respondent, or "Nursing Home Information." Finally, the fourth composite contains a variety of items that speak to the general quality of the care delivered by the nursing home. For convenience we call it "Nursing Home Care Quality," although the content refers to many specific aspects of care. See Table 3 on page 19 for the list of composites and related items.

A series of analyses were conducted to determine the measurement properties of the items using the US data. Results are reported with the US data because the survey is most likely to be used in the US; however, similar results were obtained with the Canadian data. Table 7 summarizes the disposition of the 31 items tested including the composite structure of the 21 items that were retained along with their measurement characteristics. The last column of the table explains why two of the items were kept despite having marginal measurement characteristic and why each of the 10 items was dropped.

As shown in the table, all four composites demonstrate sufficient criterion validity, as evidenced by their relatively high correlations (> 0.30) with the three global measures. Although the observed facility-level reliability of the first composite – Meeting Basic Needs – is not as high as we would like, it will be able to discriminate across nursing homes, given a sufficient number of respondents per facility. As mentioned above, this composite has a high percentage of missing data because it is made up of three items that were appropriately skipped by a large number of respondents.

Table 7. Final Composites and Dispositions

		Correlation with Globals (> .30)			NHr[c]	
Q#	**Composite or Item Text**	**Ever Unhappy**	**Recom-mend**	**0-to-10 NH Rating**	**> .70**	**Justification for Retaining/Dropping Item**
	Meeting Basic Needs: Help with eating, drinking, toileting[a]	**0.44**	**0.46**	**0.54**	**0.48**	
19	Wait too long for help with eating	0.40	0.44	0.50	0.46	
21	Wait too long for help with drinking	0.33	0.36	0.41	0.32	
23	Wait too long for help with toileting	0.57	0.51	0.56	0.55	
	Nurses/Aides' Kindness/ Respect Towards Resident	**0.52**	**0.59**	**0.70**	**0.83**	
12	Nurses/Aides treat resident with respect	0.36	0.47	0.55	0.77	
13	Nurses/Aides treat resident with kindness	0.43	0.48	0.59	0.81	
14	Nurses/Aides really cared about. resident	0.49	0.55	0.70	0.82	
15	Nurses/Aides rude to resident	0.39	0.38	0.41	0.64	
26	Nurses/Aides appropriate with violent resident	0.41	0.50	0.62	0.16	
	Nursing Home Provides Info/ Encourages Respondent Involvement	**0.50**	**0.64**	**0.66**	**0.85**	
29	Nurses/Aides give respondent information about resident	0.48	0.54	0.60	0.85	
30	Nurses/Aides explain things to respondent	0.45	0.53	0.56	0.78	
31	Nurses/Aides discourage respondents questions	0.19	0.26	0.24	0.48	Marginal measurement characteristics. But Focus Groups and Ombudsmen liked; Scaled well with its composite.
43	Respondent stops self from complaining	c	0.43	0.38	0.50	
45	Respondent involved in decisions about care	0.33	0.44	0.48	0.75	
53	Respondent given Info. about payments/expenses	0.32	0.57	0.51	0.63	
	Nursing Home Staffing, Care of Belongings, & Cleanliness	**0.55**	**0.67**	**0.77**	**0.89**	
11	Can find a nurse or aide	0.40	0.54	0.57	0.76	
51	Enough nurses/ aides?	0.46	0.57	0.64	0.88	
32	Room looks/smells clean	0.41	0.49	0.57	0.82	
24	Resident looks/ smells clean	0.42	0.44	0.55	0.85	
35	Public areas look/smell clean	0.37	0.52	0.58	0.85	

Table 7. Final Composites and Dispositions

		Correlation with Globals (> .30)			NHr[c]	
Q#	**Composite or Item Text**	**Ever Unhappy**	**Recom-mend**	**0-to-10 NH Rating**	**> .70**	**Justification for Retaining/Dropping Item**
37	FM's medical belongings lost	0.21	0.22	0.24	0.21	Marginal measurement characteristics. But Focus Groups and Ombudsmen liked; Scaled well with its composite.
39	FM's clothes lost	0.29	0.35	0.40	0.82	
Items Dropped from Composites						
16	Nurses/Aides treat resident rough	0.24	0.28	0.32	0.14	Marginal measurement characteristics. Group agreed to drop.
17	Another resident rude to respondent's family member	0.20	0.21	0.26	0.08	Marginal measurement characteristics. Group agreed to drop.
27	Nurses/Aides treat respondent with respect	0.41	0.51	0.54	0.74	Not uniquely and strongly related to its composite. Other items tap concept. Group agreed to drop.
33	Noise level around room	0.27	0.38	0.41	0.79	Not uniquely and strongly related to its composite. Group agreed to drop.
34	Places to talk to resident in private	0.26	0.31	0.35	0.76	Not uniquely and strongly related to its composite. Other items tap concept. Group agreed to drop.
36	Nurses/Aides didn't protect resident's modesty	0.20	0.23	0.23	0.56	Marginal measurement characteristics. Group agreed to drop.
42	Staff handled respondent's concerns satisfactorily	c	0.67	0.68	0.68	Includes term "satisfaction" which is not consistent with CAHPS methods; also low NH-reliability
46	Nursing Home has care conferences	0.00	-0.20	-0.15	0.82	Care conferences are required in Canada ONLY.
48	Management handled complaints satisfactorily	0.41	0.61	0.62	0.59	Includes term, "satisfaction" which is not consistent with CAHPS methods; also low NH-reliability
58	Staff did not give respondent info. due to privacy laws	-0.02	0.00	0.01	0.0	Marginal measurement properties. Group agreed to drop.

[a] Composite labels are in bold font. Composite scores are calculated as the mean the scores of the items that make up the composite. Composite score was only calculated for respondents who had non-missing data for half or more of the items that make up the composite.

[b] Nursing Home Reliability, or Inter-unit reliability (IUR). This statistic represents a transformation of the F-statistic for testing differences among nursing homes on an item or composite (IUR = (F - 1)/F). IUR can be interpreted as the fraction of the variation among nursing home scores that is due to real differences rather than chance. The higher the IUR, the greater the ability of the item or composite to discriminate across nursing homes. An IUR > 0.70 is considered to indicate a high level of discriminant ability for an item or composite.

[c] Q40 is the screener item that controls whether or not the respondent skips Q42 and Q43. Only those who responded 'yes' to Q40 were eligible to respond to either Q42 or Q43

Table 8 presents the composite-level psychometrics for the four final composites in the US data. The internal consistency reliability (Cronbach's alpha) was relatively high for all four composites, which indicates that the scores would provide reliable data. With the exception of the Nursing Home Care Quality composite, the item-total correlations and scaling success of the composites were all high. With regard to the Nursing Home Care Quality composite, these results taken together indicate that the composite has reliable scores, but overlaps in meaning with some content in the other composites. However, this finding is to be expected given that the composite is a general indicator of nursing home quality and actually indicative of the composite's validity as a more general measure. The Nursing Home Information and Nursing Home Care Quality composites both contain one item each that was weakly related to the overall total score, corrected for overlap (data not reported but available upon request). Both of these items were retained in the survey in response to the concerns of potential respondents and ombudsman and are the same items flagged in Table 7 (i.e. Q31 and Q37).

Table 8. Composite-level Psychometrics

Composite	Mean (SD)	Alpha	Median Correlation with Own Composite[a]	% Missing Data[b]	Scaling Success (%)	% at Floor	% at Ceiling
Meeting Basic Needs: Help with eating, drinking, toileting	73.5 (39.3)	0.90	.80	1.8%	100%	19%	64%
Nurses and Aides' Kindness and Respect Towards Family Members	84.8 (19.9)	0.88	.77	3.1%	100%	0%	9%
How Well the NH Provides Info and Encourages Family Involvement	87.4 (17.0)	0.78	.58	9.1%	100%	0%	15%
NH Staffing, Care of Belongings, and Cleanliness	80.5 (17.1)	0.79	.58	4.7%	86%	0%	9%

[a] Correlation between item and the composite to which it belongs, corrected for overlap. Within each composite, these correlations are indicators of convergent validity, and should, as a rule of thumb, be greater than 0.40.
[b] % Missing denotes the percentage of eligible respondents for whom the composite could not be scored without imputing missing data for the items within the composite.

The Meeting Basic Needs composite was limited in the amount of information it could provide because of limited degrees of freedom. That is, due to the dichotomous response format, there were only 5 possible scores. Thus, more than 60% of eligible respondents provided answers indicative of the highest quality care, while another 20% provided answers indicative of the

lowest level of quality of care on that composite. Variability for this composite could be improved by including more items. However, we decided not to add to the composite at this time in order to minimize respondent burden and because some stakeholders argued against the inclusion of the composite because they perceived the content as too critical of nursing homes. The other three composites demonstrated a great deal of variability in scores – very few of the scores observed were either the highest possible (at the ceiling) or lowest possible (at the floor).

Overall all four composites demonstrated good psychometric properties, as shown in the Table 9. The Nursing Home Care Quality composite lacks some discriminant validity due to its relationship to other composites. The basic needs composite has relatively low facility-level reliability, but that can be partially remedied with larger sample sizes.

Table 9. Summary of Psychometric Criteria

	Met Criterion?			
Statistical Criterion	**Basic Needs**	**Staff Kindness and Respect**	**NH Info and Encouragement**	**NH Staffing, Care, Cleanliness**
NH-Level Reliability (> 0.70)	No	Yes	Yes	Yes
Internal Consistency Reliability (> 0.70)	Yes	Yes	Yes	Yes
Criterion Validity (> 0.30)	Yes	Yes	Yes	Yes
Convergent Validity (> 0.40)	Yes	Yes	Yes	Yes
Scaling Success (~100%)	Yes	Yes	Yes	No

Case-mix Analysis

Background

In order to compare facilities included in this study, it is important to control for the influence of respondent characteristics on the outcome variables. Past research on health plans has shown that some types of respondents, such as older respondents, or those who are in better health, tend to give more positive responses to CAHPS items than other types of respondents (O'Malley, Zaslavsky, Elliott, Zaborski, & Cleary, 2005). Conversely, those respondents with more education or poorer health tend to give less positive responses to CAHPS items. These are characteristics of the respondents that are related to the CAHPS scores, but do not constitute characteristics of the facility, system, or plan being evaluated. Nor are they believed to reflect true differences in the quality of care delivered.

Generally speaking, when comparing facilities, the differences should derive entirely from differences in the quality of care they provide. The family survey items are designed to achieve this goal. However, if facility differences derive in part from differences in the respondent or resident populations in those facilities, rather than entirely from differences in the quality of care they deliver, it is important to remove the portion of the scores that come from respondent/resident characteristics so that comparisons accurately reflect differences in quality. These differences in the populations are called *case-mix* and removing their contribution to the scores is called *case-mix adjustment*.

The three basic goals of case-mix adjustment in the analysis of patient assessments of care are to:

1. Help remove the effects of individual patient characteristics that can affect patient or family member experiences,
2. Remove effects that might be considered spurious (i.e., that reflect something other than quality of care), and
3. Remove incentives for facilities to avoid enrollment of "hard to treat" patients (Zaslavsky, 1998).

Zaslavsky (1998, p.58) outlines three conditions to be met in the selection of variables for case-mix adjustment.

1. Within the facilities, the case-mix variables must be related to the outcome measures (ratings). That is, the variables must have sufficient *predictive power* in relation to the outcomes (e.g., older respondents give higher ratings of the nursing home care).

2. There must be variation between facilities on the case-mix variables. That is, the variables must be unevenly distributed across facilities (e.g., one facility might have a significantly higher percentage of elderly family respondents than another). This condition is the *heterogeneity* of the predictor.

3. The case-mix variables must be appropriate for adjustment because they are not themselves determined by the facilities' actions. That is, they must be characteristics that are brought to the facility by the respondent or patient (e.g., age or education), not characteristics that might be consequences of satisfaction with or assessment of the facility. For example, the length of relationship with the facility can reasonably be considered a *consequence*, rather than a *cause*, of a respondent's satisfaction with the nursing home care experiences of their family member.

This study is further complicated by the fact that the respondents are not the ones who are directly receiving the care, and are being asked to report their observations and experiences regarding the care received by someone else – the family member who is the actual resident in a nursing home facility. The initial choice of potential case-mix adjusters reflects an effort to take into consideration both the characteristics of the respondent (e.g., the respondent's age, gender, education, and number of times the respondent visits the resident) and the resident. One measure typically used as an adjuster for CAHPS analysis is the respondent's self-rating of their overall health. We have no such measure of the respondent's health on this survey; but we do have analogous measures of the resident's health (e.g., question items about the resident's capacity to make their own decisions or the resident's memory problems).

The case-mix analysis proceeds through the following steps:

1. Selection of potential case-mix adjusters
2. Estimation of predictive power of the selected adjusters
3. Estimation of heterogeneity
4. Calculation of explanatory power and impact of each adjuster

Selection of potential adjusters

We chose an initial set of adjusters to evaluate based on both historical use – some variables, such as age and education, have been subjected to extensive case-mix analysis in other CAHPS studies, and there is ample evidence that they are important case-mix adjusters – and the conceptual appropriateness of the variable as an adjuster. The next step was to select a subset of these potential case-mix adjusters for further analysis. The strength of the relationship of each potential adjuster to overall ratings of nursing home care was evaluated using step-wise regression in which each potential adjuster was regressed onto three global-type outcomes.[4] Our initial pool of potential adjusters consisted of the following:

- Respondent gender
- Respondent age
- Respondent education
- Resident has serious memory problems (q07)
- Resident ever had to share a room (q06)
- Expect resident to live in this or other Nursing Home permanently (q05)
- How often resident capable of making decisions about own daily life (q08)
- Number times in last 6 months respondent visited resident (q09) [5]

[4] The stepwise method is a modification of the forward-selection technique and differs in that variables already in the model do not necessarily stay there. As in the forward-selection method, variables are added one-by-one to the model, and the F-statistic for a variable to be added must be significant at the inclusion p-level. After a variable is added, however, the stepwise method looks at all the variables already included in the model and deletes any variable that does not produce an F-statistic significant at the exclusion (retention) p-level. Only after this check is made and the necessary deletions accomplished can another variable be added to the model. The stepwise process ends when none of the variables outside the model has an F statistic significant at the inclusion p-level and every variable in the model is significant at the retention p-level, or when the variable to be added to the model is the one just deleted from it. Following O'Malley et al. we set the inclusion p-value criterion at 0.10, and the exclusion (retention) criterion at 0.05 (i.e., to stay in the model, a given variable had to be significant at $p. < .05$).

[5] Respondent's age, education, q08, and q09 were treated as continuous variables. Respondent gender, q06, and q07 were dummy-coded and entered into the model as indicator variables with a value of 1 or 0. For gender, the reference category is female; for q06 and q07, which are yes/no questions, the reference category is 'no.' Since a substantial number of respondents answered "don't know" to q05, we retained that response category and dummy coded q05 – the reference category is 'no.'

We did not use relationship items relating to financing or responsibility (Q54-57) because other items reflected the interaction better, such as visiting the resident, because of the need to decrease the size of the survey and because two of the items, Q56 and Q57 which ask about legal guardianship presented problems with missing data and skip errors.

We modeled the relationship of our potential adjusters to three global outcome measures. Thus, three stepwise regression models were calculated in which the potential adjusters listed above were regressed on each of the three following outcome variables:

- Ever Unhappy with Care your Family Member Received[6] (q40)?
- Global Rating of Care at Nursing Home (q49)
- Would You Recommend this Nursing Home (q50)?

The results of the variable-selection analysis are reported in Table 10; bolded variables met the inclusion criteria for at least one of the three outcomes.

Table 10. Case-mix Adjuster Selection

	Ever Unhappy with Care[6] (q40)	Rate Care at NH (q49)	Would Recommend NH (q50)
Potential Adjusters	*Parameter Estimate*		
Respondent Gender (Male)	0.082*	0.392*	0.124[b]
Respondent Age	0.045**	0.237***	0.077**
Respondent Education	a	a	0.082**
Resident Memory Problems (q07)	a	a	a
Resident Share a Room (q06)	a	a	a
Resident In NH Permanently Yes (q05)	a	1.167**	0.415**
Resident In NH Permanently Don't Know (q05)	a	0.821[b]	a
Resident Capable of Making Decisions (q08)	0.06***	0.131*	a
Number Times Respondent Visited Resident (q09)	-0.07***	a	a

[a] Failed to meet inclusion p-value criterion.
[b] Met inclusion criterion, but failed to meet model retention criterion.
* = p < 0.05; ** = p < 0.01; *** = p < 0.001.

[6] It should be noted that q40 is coded so that a score with a quantitative value of one corresponds to a response equivalent to "I was **unhappy** with the care my family member received", and a score with a quantitative value of two corresponds to a response equivalent to "I was **never unhappy** with the care my family member received," so a higher score represents a more positive experience. The question label in the table has been changed to reflect that the more positive response has a higher quantitative value.

Predictive Power

The next step to identifying case-mix adjusters was to evaluate each of the remaining six variables for their relative unique strength in predicting the three outcomes. Predictive power was measured as the incremental amount of variance in each outcome explained by a potential case-mix adjuster (the predictor) in a step-wise linear regression analysis, controlling for facility and the other variables being assessed as potential case-mix adjusters. Following O'Malley, Zaslavsky, Elliott, Zaborski, & Cleary (2005), we report the power of each adjuster to predict this quality rating as the partial r^2 for that adjuster * 1,000. These values are presented in the Table 10.

Case-mix Adjuster Heterogeneity

Each potential case-mix adjuster was also evaluated for how much it differed between facilities compared to how much it differed within each facility. Heterogeneity of each potential case-mix variable (predictor) across plans was defined as the ratio of its between-facility variance to its residual within-facility variance. For case-mix adjuster heterogeneity, we ran a separate variance component model with each potential case-mix adjuster as the outcome, with the other potential case-mix adjusters specified as fixed effects, and facility specified as a random effect (i.e., we estimated the contribution of facility to the variance of the potential case-mix adjusters). These results are presented in Table 11 as well.

Table 11. Assessing the Impact of Selected Case-mix Adjusters

		Ever Unhappy with Care (q40)			Rate Care at NH (q49)			Would Recommend NH (q50)		
Outcome Heterogeneity Factor		***0.0843***			***0.2911***			***0.2763***		
Potential Adjusters	*Case-Mix Adjuster Heterogeneity Factor*	*Predictive Power*	*EP**	*Impact Factor**	*Predictive Power*	*EP**	*Impact Factor**	*Predictive Power*	*EP**	*Impact Factor**
Respondent Gender (Male)	-0.0018	6.80	-0.01	-0.15	11.00	-0.02	-0.07	8.50	-0.02	-0.06
Respondent Age	0.0107	11.30	**0.12**	**1.43**	21.60	0.23	0.79	8.90	0.09	0.34
Respondent Education	0.1587	4.30	**0.68**	**8.09**	2.90	**0.46**	**1.58**	0.40	0.06	0.23
Resident In NH Permanently (q05)	0.0212	4.60	**0.10**	**1.16**	16.90	**0.36**	**1.23**	13.00	**0.28**	**1.00**
Resident Capable of Making Decisions (q08)	0.0573	11.20	**0.64**	**7.60**	2.40	**0.14**	**0.47**	0.90	0.05	0.19
Number Times Respondent Visited Resident (q09)	0.0053	17.80	**0.09**	**1.11**	3.80	0.02	0.07	0.50	0.00	0.01

* Predictive Power = R-square*1,000; EP = heterogeneity * predictive power; Impact Factor = EP/Outcome heterogeneity
Bolded adjusters are those with an EP > 0.1 and an impact factor > 1.0

Impact

The first step in assessing the impact of each potential adjuster is to calculate the explanatory power (EP) of each variable being assessed – the product of its predictive power and its heterogeneity factor. Adjusters with an EP of at least 0.10 are considered to have the potential to have a noticeable impact on CAHPS scores (Zaslavsky, 1998). We then calculated the impact factor for each potential adjuster.[7]

An impact value of 1.0 for the case-mix variable indicates that it has the potential to result in a change in the outcome that is at least equal to the baseline variance in that score across plans. We used a threshold value of 1.0 for the impact factor to screen for potential case-mix variables, following O'Malley, Zaslavsky, Elliott, Zaborski, & Cleary (2005).

Recommendations for Case-mix Adjusters

We found statistical evidence supporting the possible use of five variables for case-mix adjusting (see bolded items in Table 10). Respondent age and education have historically been used as adjusters in the analysis of CAHPS data; in fact, they are recommended as adjusters in the documentation available on the CAHPS user web site. The results presented here provide empirical evidence for using them for the family survey analysis as well.

We also found support for the use of two of the health-rating analogues as case-mix adjusters. These were: the respondent's belief about resident's capability of making decisions and whether or not the respondent believes the resident will permanently live in a nursing home. Both exhibit a potential impact on two of the three outcomes, and the latter appears to have a reasonable impact on the recommendation outcome.

The final adjuster potentially measures the amount of 'data' available to the respondent in reporting their observations and experiences—one who visits more regularly would

[7] To calculate the impact factor, we examined the variance in the three outcome variables across the facilities (outcome heterogeneity) by estimating a variance component model for each of the three outcomes, with facility specified as a single random effect (i.e., we estimated the contribution of facility to the variance in the outcomes). This value is presented in the second row of Table 10.
The explanatory power is divided by the outcome-heterogeneity factor to get the impact factor – a quantitative measure of each case-mix adjusters' potential impact on the variance of the three global outcomes, standardized to the baseline variance in each of the three outcomes. If an outcome has very little baseline variance across the units of interest (facilities, in this case), an adjuster with a relatively low EP can potentially have a large impact on that outcome. Conversely, for outcomes that already exhibit substantial variance across the facilities, an adjuster would need relatively higher EP to have a noticeable impact.

presumably have more opportunities to observe what goes on in the nursing home. We could also argue that a respondent who visits more often may do so *because of* the quality of care delivered by the nursing home. For example, if the care were poor, the respondent might visit more often in order to supervise or supplement that care out of concern for the resident's health and wellbeing. Alternatively, if the care were very good, the respondent might visit more often because the experience is enjoyable. If so, frequency of visits would be endogenous with the quality scores and inappropriate as a case-mix adjuster. This possibility is enough of a concern that we propose rejecting this variable as an adjuster.

In sum, we propose that the remaining four variables be used as case-mix adjusters for the family survey: respondent age, respondent education, whether resident was permanently in the nursing home, and whether the resident was capable of making decisions.

Reporting to Nursing Homes

Based on the psychometric analyses and the case mix analysis, we developed reports with each nursing home presenting each nursing home's data and a comparison to the average. A sample report is in ***Appendix H***.

TEP review and finalization of instrument

After the analyzing the data and developing a proposed composite structure, AHRQ sent the draft composite structure to the TEP for review and held a conference call to obtain their input. The TEP had some additional questions and AIR provided input and additional analyses. Once the TEP reviewed it, the family survey was finalized and is presented in ***Appendix I, the Final Survey Instrument.***

References

Elliott, M.N., C. Edwards, J. Angeles, K. Hambarsoomians, & R.D. Hays, (2005). Patterns of unit and item nonresponse in the CAHPS Hospital survey. *Health Services Research, 40 (6), Part II:* 2096–2119.

Hurtado, M., Angeles, J., Blahut, S.A. & Hays, R., (2005). Assessment of the equivalence of the Spanish and English Versions of CAHPS® Hospital Survey. *Health Services Research, 40 (6) Part II:* 2140–2161.

Keller, S.D., O'Malley, A.J., Hays, R.D., Zaslavsky, A.M., Hepner, K.A. & Clearly, P.D. (2005). Methods used to streamline the CAHPS® Hospital Survey. *Health Services Research ,40(6) Part II*: 2057-2077.

O'Malley A.J., Zaslavsky, A.M., Hays, R.D. , Hepner, K.A., Keller, S.A. & Cleary, P.D.. (2005). Exploratory factor analyses of the CAHPS Hospital Pilot Survey responses across and within medical, surgical and obstetric Services. *Health Services Research*. *40(6), Part II*: 2078-2095.

O'Malley, A.J., Zaslavsky, A.M., Elliott, M.N., Zaborski, L. & Cleary. P.D. (2005). Case mix adjustment of the CAHPS Hospital Survey. *Health Services Research*. *40(6)*, *Part II*: 2162-2181.

Rubin, D.B. (1976). Inference and missing data. *Biometrika, 63,* 581-592.

Rubin, D.B. (1987). *Multiple imputation for nonresponse in surveys.* New York: John Wiley & Sons, Inc.

Tomarken, A.J. & Waller, N.G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. Annual Review of Clinical Psychology, (1), 31-65.

Zaslavsky, AM. (1998). Issues in case-mix adjustment of measures of the quality of health plans. *Proc Gov Soc Stat.* American Statistical Association, 56-64.

Measure #/Title/Steward

NH-028-10: Consumer Assessment of Health Providers and Systems (CAHPS®) Nursing Home Survey: Family Member Instrument (Agency for Healthcare Research and Quality (AHRQ/DHHS))

Description: The CAHPS Nursing Home Survey: Family Member Instrument is a mail survey instrument to gather information on the experiences of family members of long stay (30+ days) residents currently in nursing homes. The Centers for Medicare & Medicaid Services requested development of this questionnaire, which is intended to complement the CAHPS Nursing Home Survey: Long-Stay Resident Instrument and the Discharged resident Instrument. The Family Member Instrument asks respondents to report on their own experiences (not the resident's) with the nursing home and their perceptions of the quality of care provided to a family member living in a nursing home. The survey instrument provides nursing home level scores on 4 topics valued by patients and families: (1) Meeting Basic Needs: Help with Eating, Drinking, and Toileting; (2) Nurses/Aides' Kindness/ Respect Towards Resident; (3)Nursing Home Provides Information/Encourages Respondent Involvement; and (4) Nursing Home Staffing, Care of Belongings, and Cleanliness. In addition, the survey provides nursing home scores on 3 global items including an overall Rating of Care.

Initial In-Person Vote:

Recommended for time-limited endorsement with conditions – 14
Not recommended – 2
Abstained from vote - 4

Steering Committee Questions/Conditions for Measure Developer:	Response from Measure Developer
• In order to harmonize with the other nursing homes measures, the Steering Committee requested reconsideration of the definition of long-stay (over 100 days) and short stay (100 days or less) populations)	• The developer harmonized their definition of short and long stay residents with other NQF measures. Short-stay residents are defined as those within the home for 100 or fewer days and long term residents are those in the home for more than 100 days.
	• The denominator exclusions have been redefined in order to harmonize with the definitions used in other nursing home measures for short term residents (residents with a stay less than or equal to 100 days).