Effect of Different Imputation Methods on Factor Analyses of CAHPS Nursing Home Survey

Robert M. Baskin, Judy Sangl, Mark W. Zodet

Agency for Healthcare, Research and Quality
540 Gaither Road
Rockville, MD 20850

Disclaimer
The views expressed are those of the authors and no official endorsement by the Agency for Healthcare Research & Quality or the Department of Health and Human Services is intended or should be inferred.

Introduction

The statement on the Agency for Healthcare Research and Quality (AHRQ) website is that "The Consumer Assessment of Healthcare Providers and Systems (CAHPS) program is a public-private initiative to develop standardized surveys of patients' experiences with ambulatory and facility-level care." The word 'survey' here is synonymous with questionnaire and this study of the CAHPS Nursing Home Survey falls into the category of questionnaire validation. The CAHPS NH Surveys include three separate instruments: an in-person structured interview for long-term residents, a mail questionnaire for recently discharged short-stay residents, and a mail questionnaire for residents' family members. The specific CAHPS Nursing Home survey being considered here is the mail questionnaire for recently discharged short-stay residents.

In 2005 the in-person CAHPS Nursing Home survey underwent extensive psychometric testing for validity but the sample size of 127 respondents for the mail questionnaire for recently discharged short-stay residents was considered to be too small to support the psychometric testing. The CAHPS Expert Team has validated the use of the CAHPS Nursing Home questionnaire for use with long stay nursing home patients based on a factor analysis model with five factors. In 2009 the state of Maryland conducted a statewide survey of recently discharged short-stay residents and obtained responses from 1828 eligible responders. A project was undertaken to use the data from the 2009 Maryland survey of recently discharged short-stay residents to validate the factor analysis model from the long stay nursing home patients. Investigation of the Maryland short stay data revealed a larger amount of missingness, some due to legitimate skip patterns, than in the data from the long stay population. Therefore the research work was extended to include an evaluation of the impact of imputation on the quality of the fit of the factor analysis model.

Data for the Maryland Short Stay Nursing Home Population

As previously mentioned, the CAHPS Expert Team had validated the use of the CAHPS Nursing Home questionnaire for use with long stay nursing home patients. In 2009 the state of Maryland provided data on 1828 short stay nursing home patients who were eligible responders to the CAHPS Nursing Home questionnaire. The Maryland data came from a mailed questionnaire that was mailed to all recently discharged patients in selected Maryland nursing homes. The nursing homes were selected if they had 100 discharged patients in the prior year. There were 36 nursing homes selected with approximately 4,000 mailed questionnaires and there was a return rate on the questionnaires of 48%. Of the 1,859 returned questionnaires, 31 were deemed ineligible by CAHPS rules leaving 1,828 eligible respondents. The questionnaire for the short stay population has some features different from the long stay population. The questionnaire for the long stay population was administered in person by a trained data collector and in the CAHPS study achieved a 69% response rate. Furthermore, there are skip patterns in the short stay data that may affect the quality of the model fit.

Factor Models for the Maryland Short Stay Nursing Home Population

The CAHPS Expert team constructed a five factor model to fit the in-person structured interview administered to the long stay residents. The five factors included Environment, Care, Community, Activity, and Autonomy. The Environment factor is based on nine items related to the environment of the nursing home such as quality of the food, ability to reach items in the room, and level of noise in the room. The Care factor is based on five items relating to the patients care. The Communication factor is based on three items relating to how the nursing home staff communicates with the patient. The Autonomy factor is based on three items relating to how much autonomy a patient has. The Activity factor is based on two items relating to activities at the nursing home. The model based on these five factors is referred to as the five-factor model. The CAHPS Expert Team also believed the Autonomy factor might not be important for the Short Stay population. Typically short stay patients are admitted into Skilled Nursing Facilities with a fixed term plan for rehabilitation. The Expert Team believed that a short stay patient would be focused on rehabilitation in order to exit in the planned term and that Autonomy might be less important than for a patient with no plan for exit. The four-factor model includes Environment, Care, Communication, and Activity but eliminates Autonomy. Under the best imputation scenario, the multiple hotdeck imputation, neither the five-factor model nor the four-factor model can reject a very good fit of the factor model using the criterion of Root Means Square Error of Approximation (RMSEA). The five-factor RMSEA is 0.051 with a 95% confidence interval of (0.04755, 0.054471). Similarly for the four-factor model the RMSEA is 0.052 with a 95% confidence interval of (0.047504, 0.055613). Unadjusted Chi-squared statistics are different for the two models because of very different degrees of freedom but the adjusted statistics are generally similar. Note that the two Bayesian Information Criterion numbers are not directly comparable because of the absence of questions 33 to 35 in the four-factor model. The output from these models, both using multiple hotdeck, is in Appendix A.

Cronbach's Alpha for the Five Factors

Cronbach's alpha was calculated for the five factors using both multiple hotdeck imputation and mice (multiple imputation using chained equations). Alphas were also calculated for subgroups determined by the two skip patterns. Overall the alphas were similar between the multiple hotdeck imputations and the mice imputations. For the overall group, the four factors for Environment, Care, Community and Activity were generally above 0.75 under mice imputation and above 0.70 for the hotdeck imputation. However, the alpha for the factor Autonomy was 0.58 under mice and 0.56 under multiple hotdeck imputation. Usually a Cronbach alpha this low is considered indicative of a problem with reliability of the measurement of the factor. The amount of missingness in the three questions Q33, Q34, and Q35, which load onto the Autonomy factor was 1.9%, 1.4% and 5.3% respectively. This is a relatively small amount of missingness and it is doubtful that the missingness or the form of imputation contributes to the low alpha.

Amount of Missingness

The Maryland short stay questionnaire included two skip patterns for items in the factors for Environment and Care. No current software for factor analysis can retain missing values in the items without using either implicit or explicit imputation so if the legitimately skipped questions are to be retained in the analysis the skipped values must be imputed, either implicitly or explicitly since subsetting to one third of the data is not an option. For this study a single hotdeck, multiple hotdeck, multiple imputations with chained equations (mice), and multiple imputation under Proc MI in SAS were all used.

Question three, relating to the dining hall, contributed to the Environment factor, and for question three approximately two thirds of the respondents had legitimate skip patterns. For all the remaining items in the Environment factor, namely simplified reaching and questions one, four, five, six, twenty, twenty one and twenty three the missingness was less than six percent. Question eight, which contributed to the Care factor, had approximately twenty percent legitimate skips and a total of 21.6% missing. In addition, three other items in the Care factor, questions nine, twelve, and thirty two, had over twenty percent missing. Question twelve was the final item contributing to the Care factor and it had 2.4% missing. Questions thirteen, fourteen, and fifteen, which composed the Communication factor, each had two percent missingness or less. The Autonomy factor is comprised of questions thirty three, thirty four, and thirty five with 1.9%, 1.4%, and 5.3% missingness respectively. Questions

thirty six and thirty seven were the only items in the Activity factor and these items had 12% and 11% missingness respectively. A complete list of items with missing percent is given in appendix B.

Method of Imputation

For each of these methods, single hotdeck, multiple hotdeck, multiple imputations with chained equations (mice), and multiple imputation under Proc MI, missing values were imputed and a confirmatory factor analysis based on the Long Stay five-factor model was run. Previously in the Long Stay study a single imputation from Proc MI was used. The Proc MI imputation is based on the assumption that the variable being imputed is normal and a draw is made from a normal distribution. This may provide imputations that work in factor analysis but the individual imputations can produce invalid values for the question being imputed. For example, Proc MI imputed a variable on an ordinal scale of zero to ten to have a value of -17. This may not be noticeable in a factor analysis on the full dataset but this imputation produces ludicrous results at the nursing home level for summary statistics.

The single hotdeck was originally run to quickly provide valid imputation values and assess the factor model under a valid imputation. However, single hotdeck imputation does not properly account for the variance due to imputing. Later, mice was run to provide valid imputations and account for the variance in the imputation. However, there appeared to be a difference in the fit of the five-factor model under the two imputations so in addition, a multiple hotdeck was run. The multiple hotdeck does produce valid imputations, the five-factor model does have a similar model fit under multiple hotdeck as under single hotdeck, and it does allow hypothesis tests to include variability due to imputation.

Confidence intervals for the RMSEA statistics provide a method to compare the models under mice and multiple hotdeck. For the five-factor model fit to the data with multiple hotdeck imputations, the RMSEA index is 0.050786 with a 95% confidence interval of (0.047343, 0.054267). For the five-factor model fit to the data using mice imputations, the RMSEA index is 0.067611 with a 95% confidence interval of (0.06424, 0.071017). Because the intervals are not overlapping, this is equivalent to rejecting a conservative 90% test that the RMSEA for the two are the same. Note that the models are the same and the observed data is the same and the only difference is the method of imputation.

The investigation of the results indicated that Proc MI, in this situation, is problematic for multiple reasons. The obvious problem of imputing invalid responses was mentioned and in fact some of the responses were imputed by Proc MI to be far out of range. Some experts claim that this does not affect the factor analysis but there is some indication from the confirmatory factor analysis that the imputation from Proc MI does create problems for the factor analysis. When performing factor analysis on the Proc MI imputed data in both R and SAS, error and warning messages were produced that indicated there is some problem with the Hessian matrix. Since the factor models and observed data are the same as with the mice and multiple hotdeck imputed data and since these imputations did not produce the same errors and warnings it must be concluded that there is some problem with the Proc MI imputations that is affecting the factor analysis.

Subset Analysis

For subsets determined by questions two and seven, the questions determining skip patterns, factor analysis models were run on the subsets and Cronbach's alpha was calculated for each subset for the imputation methods of multiple hotdeck and mice.

The output for the subset factor analysis is in the complete output provided but it is quite lengthy. However, neither type of imputation showed much more than random variation across the subsets in the fit statistics. The sample size for the two subsetting questions was unbalanced with 33% observed data for question three and 22% imputed data for question eight. This might possibly make it more difficult to detect differences.

The story for the Cronbach's alpha was quite different. For the subsets determined by question two, there were approximately 66% imputed data for question three with legitimate skips. Intuition might suggest that if there were any problem with the imputations it would show up in this group, but the Cronbach's alpha across the subsets for the Environment factor, which incorporated question three, showed not much more than random variation. For the

subsets determined by question seven there were approximately 21% with legitimate skips. Looking at the Care factor, which incorporated question eight, the mice imputed data produced alpha values of 0.84 (to two decimal places) for the overall data and for both subsets. However, the alpha values for the Care factor in the two subsets determined by question seven and using the multiple hotdeck data were 0.80 for the observed data and 0.39 for the data with legitimate skips. The interpretation of this is that in the legitimate skips of question seven the mice imputations are retaining whatever structure is in the observed data but the multiple hotdeck is imputing question eight in a way that may make the marginal distribution look valid but is not imputing the question in a way the structure for the Care factor hangs together.

Conclusion

The method used was to imitate the factor analysis from the long stay study as closely as possible. The analysis indicates that the method of imputation has an effect on the fit of the factor model. The multiple hotdeck might seem to have an advantage in that regard. However, the subset analysis indicates that for at least one factor the multiple hotdeck is not imputing in a way that retains the structure of the observed data. Perhaps the safest approach is to give more credence to the mice imputations and accept that the fit is not in the very good range but is plausible. Unfortunately the results indicate the quality of the questions does not allow a determination if the five or four-factor model is more appropriate.

References

Carol Cosenza, Floyd J. Fowler, Jr., Joan L. Buchanan, Paul D. Cleary, and Lin Ding
Nursing Home CAHPS Field Test Report , January 19, 2006

Appendix A

These models are based on the multiple hotdeck imputation.

Output from the Five Factor Model:

```
 Model Chisquare =  1115.6   Df =  194 Pr(>Chisq) = 0
 Chisquare (null model) =  16316   Df =  231
 Goodness-of-fit index =  0.94537
 Adjusted goodness-of-fit index =  0.92875
 RMSEA index =  0.050991   95% CI: (0.04755, 0.054471)
 Bentler-Bonnett NFI =  0.93162
 Tucker-Lewis NNFI =  0.93178
 Bentler CFI =  0.9427
 SRMR =  0.039808
 BIC =  -341.55

 Normalized Residuals
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-2.690  -0.656   0.000   0.290   0.723   9.070

Iterations =  167
mod.indices(cfaMHDout)

 5 largest modification indices, A matrix:
 Q.09:Q.08  Q.20:Q.21  Q.21:Q.20 Reach:Q.33  Q.08:Q.09
 124.28184  100.09846  100.08827   99.58666   88.94117

  5 largest modification indices, P matrix:
 Q.09:Q.08  Q.21:Q.20   F4:Reach Q.33:Reach    F5:Q.35
 195.90698  119.17329  111.92599   60.50942   50.71777
```

Output from the Four Factor Model:

```
 Model Chisquare =  825.1   Df =  141 Pr(>Chisq) = 0
 Chisquare (null model) =  15108   Df =  171
 Goodness-of-fit index =  0.9532
 Adjusted goodness-of-fit index =  0.93694
 RMSEA index =  0.051532   95% CI: (0.047504, 0.055613)
 Bentler-Bonnett NFI =  0.94539
 Tucker-Lewis NNFI =  0.94446
 Bentler CFI =  0.9542
 SRMR =  0.034987
 BIC =  -233.95

 Normalized Residuals
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-2.5600 -0.5880  0.0001  0.2380  0.6820  9.0800

 5 largest modification indices, A matrix:
Q.09:Q.08 Q.20:Q.21 Q.21:Q.20 Q.08:Q.09   F4:Q.23
124.51297 100.33191 100.07383  88.86936  43.22949

  5 largest modification indices, P matrix:
 Q.09:Q.08  Q.21:Q.20    F4:Q.23  Q.10:Q.08 Q.32:Reach
 196.29318  119.29343   50.13654   25.21022   23.67360
```

Appendix B

For Each Item in a Factor the Missingness is Presented as a Proportion

Environment Factor

reachsi2	q01	q03x	q04	q05	q06	q20rf	q21r	q23xrf
0.0383	0.0334	0.6603	0.0213	0.0159	0.0164	0.0148	0.0098	0.0531

Care Factor

q08x	q09x	q10	q12x	q32xrf
0.216	0.222	0.024	0.253	0.269

Communication Factor

q13	q14	q15
0.011	0.014	0.020

Autonomy Factor

q33rf	q34rf	q35rf
0.019	0.014	0.053

Activitiy Factor

q36rf	q37rf
0.12	0.11