Health Utilities Index Data Derived from the 2000 Health and Retirement Study

Duncan Ermini Leaf
Leonard D. Schaeffer Center for Health Policy and Economics
University of Southern California
dleaf@healthpolicy.usc.edu

February 27, 2015

1 Background

The Health Utilities Index Mark 3 (HUI3) is a system for health status classification and health related quality-of-life (HRQoL) scoring. Horsman et al. (2003) provide an overview of HUI systems with supporting references. The 33-item HUI3 questionnaire (HUI3SU.33Q) was included as experimental module 7 in the 2000 Health and Retirement Study (HRS) wave. The Schaffer Center for Health Policy Economics obtained the HUI3 scoring procedures (Furlong et al., 2001) from HUInc. and applied them to the questionnaire responses in the 2000 HRS. The derived HUI3 variables are provided as a contribution to the HRS research community. This file documents the data file of derived variables and accompanying imputation files.

2 Data description

HUI3 data were gathered from 1,156 respondents in the 2000 HRS wave. Table 1 lists the variables included in derived HUI3 data set. The variables fall into three categories: HRS identifiers that link the HUI3 derived data to HRS respondents, HUI3 attribute levels that define a respondent's health state, and HUI3 utility scores that represent mean preference scores for the general population. The codebook for these data is included as a separate file.

2.1 HRS identifiers

The *hhid*, *pn*, and *hhidpn* variables can be used for merging derived HUI3 scores with the HRS or RAND-HRS data sets.

Variable name	**Type**	**Description**
hhid	character	HRS household ID
pn	character	HRS person number
hhidpn	numeric	hhid concatenated with pn
hui3vl	numeric	HUI3 Vision level
hui3hl	numeric	HUI3 Hearing level
hui3sl	numeric	HUI3 Speech level
hui3al	numeric	HUI3 Ambulation level
hui3dl	numeric	HUI3 Dexterity level
hui3el	numeric	HUI3 Emotion level
hui3cl	numeric	HUI3 Cognition level
hui3pl	numeric	HUI3 Pain level
hui3vu	numeric	HUI3 Vision utility score
hui3hu	numeric	HUI3 Hearing utility score
hui3su	numeric	HUI3 Speech utility score
hui3au	numeric	HUI3 Ambulation utility score
hui3du	numeric	HUI3 Dexterity utility score
hui3eu	numeric	HUI3 Emotion utility score
hui3cu	numeric	HUI3 Cognition utility score
hui3pu	numeric	HUI3 Pain utility score
hui3ou	numeric	HUI3 Overall Utility Score
hui3os	character	HUI3 Overall Health State

Table 1: Contents of the HUI3 data set derived from module 7 in the 2000 HRS

2.2 HUI3 attribute levels and overall health state

The attribute level variables, *hui3vl*, *hui3hl*, *hui3sl*, *hui3al*, *hui3dl*, *hui3el*, *hui3cl*, and *hui3pl*, were derived from the 2000 HRS module 7 responses according to instructions in the HUI3 manual (Furlong et al., 2001). More information about these attribute levels can be found in Feeny et al. (1995, 1996, 2002) and Furlong et al. (1998).

The overall health state (*hui3os*) is a vector of the attribute levels stored as a string. The format is: (*hui3vl*,*hui3hl*,*hui3sl*,*hui3al*,*hui3dl*,*hui3el*,*hui3cl*,*hui3pl*). For example, if a respondents has *hui3vl*=1, *hui3hl*=1, *hui3sl*=2, *hui3al*=1, *hui3dl*=3, *hui3el*=3, *hui3cl*=1, and *hui3pl*=2, then *hui3os* = "(1,1,2,1,3,3,1,2)" for that respondent. Each value of *hui3os* is a unique combination of the attribute levels.

2.3 HUI3 utility scores

The mapping from attribute levels to the HUI3 single-attribute utility scores, *hui3vu*, *hui3hu*, *hui3su*, *hui3au*, *hui3du*, *hui3eu*, *hui3cu*, and *hui3pu*, can be found in Furlong et al. (1998) and Feeny et al. (2002). Each single-attribute utility score is defined to be in the range from 0 (most morbidity) to 1 (no morbidity).

The formula for computing the HUI3 multi-attribute utility score (*hui3ou*) from the attribute levels can also be found in Furlong et al. (1998) and Feeny et al. (2002). The overall utility score is a health-related quality of life measure. It is defined to be in the range from -0.36 to 1. A score of 0 is equivalent to death, 1 is perfect health, and negative scores represent health states considered worse than death.

3 Imputation of missing attribute levels

The HUI3 questionnaire has a block of questions for each attribute. For some HRS respondents, an exact HUI3 attribute level could not be derived because one or more responses in the block of questions was missing or incomplete. Without an attribute level, it is impossible to compute the single-attribute utility score, overall utility score, and the health state vector. The number of missing cases for each attribute level can be found in the accompanying codebook. The *Hearing* attribute level is most problematic with 66 cases missing. The overall utility score and overall health state have 90 missing cases, leaving 1,066 nonmissing cases.

Many researchers will wish to apply some imputation strategy to the missing values. While the HUI3 manual discusses options for handling missing data, it does not provide any specific rules for imputation (Furlong et al., 2001). Therefore, only the data set described in Section 2 is valid according to the HUI3 derivation rules. Two sets of partial imputations are included with the valid data set in order to assist researchers in developing their own imputation models. These imputation data sets are described in Sections 3.1 and 3.2.

3.1 Logical limits

By considering a respondent's nonmissing responses for an attribute, we find that there is a logical lower and upper limit to what their attribute level could be had they responded completely to the entire block of questions. The logical limits for each attribute level are stored in variables with the same name as the attribute level, but ending in *_l* and *_u*. The *_l* suffix denotes the lower limit and *_u* denotes the upper limit. For example, the *Vision* attribute level lower limit is stored in *hui3vl_l* and the upper limit is stored in *hui3vl_u*. The upper and lower limits are equal to the derived attribute level for nonmissing cases.

Using the lower and upper limits of the attribute level, the logical limits for the single-attribute utility score, the overall utility score, and overall health state were also derived. The single-attribute utility score limits were computed using each *_l* and *_u* attribute level limit variables and then stored in variables with the same name as the single-attribute utility score, but ending with the corresponding *_l* or *_u* suffix. For example, *hui3vu_l* was computed from *hui3vl_l* while *hui3vu_u* was computed from *hui3vl_u*, both using the mapping in Feeny et al. (2002). The overall utility score limits were computed from the corresponding collection of *_l* and *_u* attribute level limit variables and then stored in the *hui3ou_l* and *hui3ou_u* variables. For example, *hui3ou_l* was computed from all of the attribute level lower limits, *hui3vl_l*, *hui3hl_l*, *hui3sl_l*, *hui3al_l*, *hui3dl_l*, *hui3el_l*, *hui3cl_l*, and *hui3pl_l*, following the formula in Feeny et al. (2002). Finally, the overall health state vector limits were derived from the corresponding collection of *_l* and *_u* attribute level limit variables and then stored in the *hui3os_l* and *hui3os_u* variables. For example, *hui3os_l* is the concatenation of all of the attribute level lower limits: *hui3vl_l*, *hui3hl_l*, *hui3sl_l*, *hui3al_l*, *hui3dl_l*, *hui3el_l*, *hui3cl_l*, and *hui3pl_l*.

The lower and upper attribute level limits are the same in some cases, resulting in a logical imputation of the exact attribute level value. These logical imputations reduced the number of missing overall utility scores and overall health states by 6 cases to a total of 1,072 nonmissing cases.

It is important to keep in mind that the *_l* and *_u* suffixes respectively identify the logical lower and upper limits of the **attribute levels**. HUI3 utility scores are typically decreasing as attribute levels increase. Therefore, utility scores ending in *_l* will typically be greater than or equal to the corresponding utility scores ending in *_u*.

All imputation models should respect the logical limits described here. An imputed value outside of these limits is inconsistent with the HUI3 system.

3.2 Conditionally-almost-sure imputations

To develop an imputation model for missing attribute levels, some researchers will use HRS data to estimate a probability distribution for the attribute level. This probability distribution could be conditional upon a respondent's observed questionnaire responses in the block of questions for the attribute. For some attributes, HRS respondents that

have the same pattern of observed responses and have a nonmissing attribute level all share the same attribute level. If we assume the attribute level is missing-at-random, then the empirical estimate is a degenerate distribution, assigning 100% probability to a single attribute level.

These cases were identified and their attribute levels were imputed at the level with 100% conditional probability. The imputed attribute levels are stored in variables with the same name as the attribute level, but ending with a *_i* suffix. For example, the imputed *Vision* level is stored in *hui3vl_i*. Additionally, when the logical lower and upper limits for the attribute level were identical (see Section 3.1), then the logical imputation was used as the imputed value. For example, *hui3vl_i* = *hui3vl_l* if *hui3vl_l* = *hui3vl_u*. The imputed value is equal to the derived attribute level for nonmissing cases.

Using imputed attribute levels, the single-attribute utility score, overall utility score, and overall health state were also derived. The imputed single-attribute utility score was computed using each *_i* imputed attribute level variable and then stored in a variable with the same name as the single-attribute utility score, but ending with the *_i* suffix. For example, *hui3vu_i* is computed from *hui3vl_i* using the mapping in Feeny et al. (2002). The imputed overall utility score, *hui3ou_i*, was computed from the imputed attribute level variables, *hui3vl_i*, *hui3hl_i*, *hui3sl_i*, *hui3al_i*, *hui3dl_i*, *hui3el_i*, *hui3cl_i*, and *hui3pl_i*, following the formula in Feeny et al. (2002). Finally, the imputed overall health state vector, *hui3os_i*, is the concatenation of all of the imputed attribute level variables: *hui3vl_i*, *hui3hl_i*, *hui3sl_i*, *hui3al_i*, *hui3dl_i*, *hui3el_i*, *hui3cl_i*, and *hui3pl_i*.

These conditionally-almost-sure imputations reduced the number of missing overall utility scores and overall health states by 66 cases to a total of 1,132 nonmissing cases.

4 Files provided

The following data files are included in this distribution. Each file is provided in SAS (.sas7bdat), Stata (.dta), R (.Rdata), and CSV (.csv) formats.

- **hrs2000_hui3_valid**: HRS identifiers and HUI3 variables derived strictly according to Furlong et al. (2001) as described in Section 2. Variables have missing values if they could not be derived because of missing or "don't know" responses in the questionnaire. No imputations of any kind are provided in this file. The list of variables can be found in Table 1.

- **hrs2000_hui3_logical_limits**: HRS identifiers and Logical lower and upper limits for the HUI3 derived variables as described in Section 3.1.

- **hrs2000_hui3_imputed**: HRS identifiers and Conditionally-almost-sure imputations described in Section 3.2.

In addition, the codebook for the HUI3 derived variables in **hrs2000_hui3_valid** can be found in the accompanying file, **hrs2000_hui3_codebook.txt**.

5 Acknowledgements

This work was supported by NINR R01NR013372.

References

Feeny, D., Furlong, W., Boyle, M., and Torrance, G. W. (1995). Multi-attribute health status classification systems. *Pharmacoeconomics*, 7(6):490–502.

Feeny, D., Furlong, W., Torrance, G. W., Goldsmith, C. H., Zhu, Z., DePauw, S., Denton, M., and Boyle, M. (2002). Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Medical care*, 40(2):113–128.

Feeny, D. H., Torrance, G. W., and Furlong, W. J. (1996). Health Utilities Index. In Spilker, B., editor, *Quality of Life and Pharmacoeconomics in Clinical Trials*, chapter 26, pages 239–252. Lippincott-Raven Press, Philadelphia, second edition.

Furlong, W., Feeny, D., Torrance, G. W., Goldsmith, C., DePauw, S., Boyle, M., Denton, M., and Zhu, Z. (1998). Multiplicative multi-attribute utility function for the Health Utilities Index Mark 3 (HUI3) system: a technical report. Technical Report 98-11, McMaster University Centre for Health Economics and Policy Analysis, Hamilton ON, Canada.

Furlong, W. J., Feeny, D. H., and Torrance, G. W. (2001). *HEALTH UTILITIES INDEX (HUI) procedures manual: algorithm for determining HUI Mark 3 (HUI3) health status classification levels, health states, health-related quality of life scores and single-attribute level utility scores from 33-item interviewer-administered health status questionnaires.* Health Utilities Inc., Dundas ON, Canada.

Horsman, J., Furlong, W., Feeny, D., and Torrance, G. (2003). The Health Utilities Index (HUI®): concepts, measurement properties and applications. *Health and Quality of Life Outcomes*, 1(1):54.