Imputing Physical and Mental Summary Scores (PCS and MCS) for the Veterans SF-12 Health Survey in the Context of Missing Data

Avron Spiro III, Ph.D.
William H. Rogers, Ph.D.
Shirley Qian, M.S.
and
Lewis E. Kazis, Sc.D.

September 28, 2004

Technical Report prepared by:
The Health Outcomes Technologies Program,
Health Services Department,
Boston University School of Public Health, Boston, MA

and

The Institute for Health Outcomes and Policy,
Center for Health Quality, Outcomes and Economic Research,
Veterans Affairs Medical Center, Bedford, MA

Supported by the Center for Medicare and Medicaid Studies through the National Center for Quality Assurance; the Health Outcomes Technologies Program, Health Services Department, Boston University School of Public Health, and the Center for Health Quality, Outcomes and Economic Research, Bedford VA Medical Center, Department of Veterans Affairs.

Questions concerning this work can be e-mailed to Drs. Spiro, Roger, Kazis or Ms. Qian at:
aspiro3@bu.edu, whrogers@comcast.net; lek@bu.edu;qian_shirley@yahoo.com

The Centers for Medicare & Medicaid Services' Office of Research, Development, and Information (ORDI) strives to make information available to all. Nevertheless, portions of our files including charts, tables, and graphics may be difficult to read using assistive technology.

Persons with disabilities experiencing problems accessing portions of any file should contact ORDI through e-mail at ORDI_508_Compliance@cms.hhs.gov

Executive Summary

This report describes a method for estimating the Physical Component Score (PCS) and the Mental Component Score (MCS) from the Veterans SF-12 Health Survey in the context of missing data. We describe a new method, modified regression estimation, for scoring observations with missing data. In addition, we present a SAS© macro implementing this method, and detail its use. Finally, we present the results of alpha testing of this version on a small sample of analysts.

Table of Contents

Attachments:
1. SAS© Programs: v-sf12-impute1.2.sas, sample.sas
2. Data: sample12.sas7bdat, pcs.sas7bdat, mcs.sas7bdat
3. Documentation (this report): sf12 imputation manual r1.doc
4. Results of alpha testing

1. Introduction to the Problem

The US Center for Medicare and Medicaid Studies (CMS) is conducting the Medicare Health Outcomes Survey (HOS) to determine the health change of Medicare beneficiaries in a variety of health plans (Cooper et al., 2001; Haffer & Brown, 2004; Jones et al., 2004). The process involves surveying beneficiaries before and after a two-year period. A similar process has taken place in the US Department of Veterans Affairs (VA) since 1996 with follow-up periods ranging from 17 months to 5 years (Kazis et al., 1998, 1999, 2000). The HOS is currently using the MOS SF-36 to conduct these surveys (Gandek et al., 2004; Ware et al., 1993, 1994); the VA has used the Veterans SF-36 and Veterans SF-12 through the Office of Quality and Performance.

It would be simple to analyze these data if everyone answered every question. However, all survey work must deal with certain practicalities related to missing data, e.g., respondents may refuse to answer some or all items in a survey. In this report, we describe one particular method for dealing with this situation, where a respondent fails to answer some items.

There are a variety of methods for dealing with the problem of missing item responses. Traditionally, cases with missing data have been omitted from the analysis (listwise deletion). Other methods include mean imputation (substituting the item mean for missing responses, which is essentially the method used in the half-scale rule adopted for scoring the MOS SF-36), regression estimation (RE), multiple imputation (MI; Little & Rubin, 1987), and the missing data estimation (MDE) method developed by Ware (e.g,. Kosinski et al., 2000; .

The method that we propose (see below), modified regression estimation (MRE), is a general method for obtaining scale scores in the context of missing data. We have developed this method in a specific context, which is the potential future use of the Veterans SF-12 Health Survey in the Medicare HOS. This method was previously used in an other report submitted to NCQA/CMS for imputation of missing values for the MOS SF-36 and Veterans SF-36 versions (Rogers et al. July 2004, Imputing the Physical and Mental Summary Scores (PCS and MCS) for the MOS SF-36 and the Veterans V/SF-36 in the presence of Missing Data). However, the method is quite general and can be applied in a wide variety of circumstances. One reason for proposing this method for this particular context is that it can be easily implemented by the end-user on a personal computer running a typical implementation of the Statistical Analysis System (SAS) software. Other methods (e.g., missing data estimation [MDE] or multiple imputation [MI]) either rely on proprietary information or are more difficult for an end-user to implement as they require higher speed computers.

Ultimately, the success or failure of any set of methods must be judged in terms of its success in any particular application. In practical terms, is the method simple to use, and can the naïve user apply it successfully? In statistical terms, is the answer invalid (biased) or imprecise? In order to understand this, we need to appeal to external data of some kind. The ultimate accuracy of the imputation method comes from its mean square error in an application, which combines bias and variance. The bias is fixed by the estimator and the nature of the comparison, but the variance depends on the sample size. A slightly biased imputation may be preferred if it can be scored in a larger sample, but this benefit is limited if the sample size is sufficiently large anyway.

A particular imputation method may be very biased in one application, but nearly unbiased in another. For example, an estimate may be biased for determining individual health status, biased for determining the physical and mental summaries from the SF-36 (PCS or MCS) associated with a disease state, but adequate for comparing health plans in the HOS or geographic service regions (VISNs) in the VA. If the purpose of estimation is general, and it does not matter whether comparisons are made with one scale or another (e.g. physical functioning or bodily pain) and these are conveying roughly the same information, then we are free to impute boldly because there is relatively little bias. However, when the exercise involves PCS and MCS comparisons between health plans, then bias may be important to identify and minimize with methods of imputation.

2. The Veterans SF-12 Health Survey

The Veterans SF-12 (Kazis et al. 1997, 1999) was developed from the Veterans SF-36 (Kazis, 2000; Kazis et al., 2000, 2004a,b), which was modified from the MOS SF-36 based on suggestions from Ware (1996). The modifications made in the Veterans SF36 are (a) an increase in the number of response choices for the role physical (RP) and role emotional (RE) items from a dichotomized two point yes/no choice to a five-point Likert scale (none of the time, a little of the time, some of the time, most of the time, all of the time), to reduce floor and ceiling effects, and (b) the use of two items to assess health change, one focusing on physical health and one on emotional problems, in contrast to the one general item in the MOS SF-36. Scoring of the Veterans SF-36 scales (Kazis et al. ,1999, 2000, 2004a, b) is similar to that for the MOS SF-36 (Ware & Kosinski, 2001; Ware, Kosinski & Keller, 1994; Ware, Snow; Kosinski & Gandek, 1993). This process includes computing scale scores if at least half of the items on a scale are present, transforming raw scores to a range from 0 to 100, where 100 denotes the best health, and computing PCS and MCS scores with a mean of 50 and a standard deviation of 10 (normed to the US population), only if scores are valid on all 8scales.

The Veterans SF-12 (Kazis et al. 1997, 1999) stands in relation to the Veterans SF-36 as the MOS SF-12 stands to the MOS SF-36 (Ware et al. 1996). It includes 1 or 2 items from each of the eight scales in the SF-36: *physical functioning, role limitations due to physical problems, bodily pain, general health perceptions, energy/vitality, social functioning, role limitations due to emotional problems and mental health*. The items were chosen on the basis of their ability to predict the component scores of the SF-36 (Ware, Kosinski & Keller, 1995, 1996).

The 12 items are used to compute a *physical component summary (PCS)* and *mental component summary (MCS).* In the Veterans SF-12, the scoring of PCS and MCS is based on weights derived from the Veterans SF-36 administered to 877,775 respondents in the 1999 Large Health Survey of Veteran Enrollees. The weights were obtained by replicating in the VA survey the method used to create the original SF-12 (Ware, Kosinski, & Keller, 1995, 1996). That is, dummy indicators were defined for response choices for each of the 12 items in the MOS SF-12, and these were then entered into multiple regressions to predict PCS and MCS scores based on the Veterans SF-36. The resulting weights, and the constant term, can be used to compute PCS and MCS scores from the Veterans SF-12 (see Appendix A).

PCS and MCS scores for the Veterans SF-12 are computed similarly to the MOS SF-12 (see Section 3 below). Compared to the MOS SF-12, the Veterans SF-12 adds about 5% more precision to the PCS and MCS. Cronbach alpha (internal consistency reliability) estimates for the Veterans SF-12 PCS and MCS are both 0.90[1].

The Veterans SF-12 has been administered in national VA surveys in 1997 and 1998 to over 60,000 patients. Since 2002, the VA has administered the Veterans SF-12 to approximately 432,000 patients annually as part of its quality management program (Survey of Health Experiences of Patients, SHEP).

3. Scoring the Veterans SF-12: Complete Data

For cases with complete data, there are three steps involved in scoring and calculating the PCS and MCS scores from the Veterans SF-12. These follow the similar approach as the MOS SF-12 (Ware et al. 1995, 1996), with some modifications for the Veterans version of the SF-12. The specifics are as follows:

Step One: Responses are first examined for out of range values (which are set to missing). Next, indicator variables are created for each response choice for each of the 12 SF-12 items, omitting one level of response. An indicator variable is not coded for the response choice category that is the lowest health state for an item (i.e., when the value of the response is 1). Taking the PF02 item as an example, there are 3 response choices, which are used to create 2 indicator variables (pf2r2, pf2r3), one indicating that response 2 was selected (pf2r2=1 if 2 was recorded, else pf2r2=0), and one indicating that response 3 was selected (pf2r3=1 if 3 was recorded, else pf2r3=0).

Note that the number of response choices for the Veterans SF-12 version differs from that of the standard MOS SF-12. This is due to the modifications to the four role limitation items (both limitations due to physical and emotional problems) where there are 5 response choices for each item instead of the 2 dichotomized choices in the original MOS SF-12 form. Of 59 total possible response choice categories among the Veterans SF-12 items, 47 indicator variables are created.

Step Two: Aggregate scores are computed, separately for PCS and MCS, by a regression equation that weights each of the 47 indicator variables (Appendix A). These weights are derived from the Veterans SF-36 Physical and Mental Component Summary Scales from the 1999 Large Health Survey of Veteran Enrollees (Kazis et al., 2000).

Step Three: The last step involves computation of the PCS and MCS by adding a constant to each of the estimates obtained in step 2. The resulting scores are set to a mean score of 50 and standard deviation of 10 for a general U.S. population as in the Veterans SF-36.

---

[1] The estimated reliability coefficients were obtained by multiplying coefficient alpha for Veterans SF-36 PCS and MCS (.96, .95), from Kazis et al., 2004a, by the explained variance of SF-36 scores by SF-12 items.

4. Scoring the Veterans SF-12: Missing Data

In the case where missing item responses are a concern, a modification has been made to Step 2 in the above approach. As discussed above, we considered several alternative approaches to estimating PCS and MCS scores for the Veterans SF-12 in the context of missing data, but for several reasons, we have adopted the modified regression estimation (MRE) approach. In addition to its advantageous statistical properties (see below), the MRE approach is also preferable because it can be implemented using a relatively simple program (described below) on a microcomputer running the base SAS system.

When there are missing item responses, we modify the regression estimation approach described in Step 2 above. For each possible combination of missing data (and for 12 items, there are 12^2 or 4096 such combinations). Thus, depending on the pattern of missing item responses, a different set of regression weights, where some are given a value of 0 for missing items, is required, one such set for each combination of missing items.

To permit estimation of PCS and MCS scores, we have estimated from the 1999 VA survey 4096 sets of coefficients for predicting SF-36 PCS and MCS scores. Each set is indexed by a variable (named "number") which runs from 0 (all 12 items present) to 4095 (all items missing), identifying the pattern of missing SF12 items. This variable "number" can be viewed as a 12-binary digit number, where a 1 means missing and a 0 means present. So 0 is no missing, 1 is the last item missing, 2 is the 2nd to last item missing, 3 is both of the last two items missing, 4 is the third to last missing, and so forth.

Separately, for each combination of missing data, the user's data are merged with the stored regression weights and PCS scores are computed and output; the process is then repeated for MCS scores. These two sets of scores are combined by a user-specified identification variable, and a new SAS dataset is created that can be saved, or merged with the user's dataset for subsequent analyses.

In practice, there should be some cutoff on our willingness to score the SF-12 with partial data. Analysis of the R-squared values for PCS and MCS show that simple rules may be inappropriate. Take PCS as an example. With 10 items present but pf02 and pf04 missing, it is possible to get R^2 above 86 percent, but if both of those items are present along with 4 other items, the R^2 can be less than 67 percent.

However, as measurement error increases, a "regression to the mean" phenomenon starts to appear. This phenomenon is related to R^2 and possibly also to the type of group being studied. The overall mean PCS in the 1999 VA sample is 36.02. Now suppose that we consider a subpopulation that has a "true" PCS of 50 but otherwise typical of persons with a PCS of 50. If the R-square is 0.9 then regression to the mean would cause the observed PCS in our sample to be (50 – 36.02) * (1-sqrt(0.9)) too low, or 0.72 points too low (i.e. 49.28 instead of 50). We could correct for this in the Veterans SF-12 computation by stretching the observed value by a factor of 1/r where r is the square root of R-squared in the equation for predicting PCS or MCS from a set of items. The full formula in this case would be:

$$\text{PCS(estimated)} = 36.02 + (\text{PCS12 Computed} - 36.02) / r$$

For MCS, this computation would be based on an observed mean of 45.39.

Note that regression to the mean was not much of a factor in the original SF-12 because the items were selected to have R-squares over 0.98[2]. In addition, if we had persons with a PCS of 50, but we found them in a pain clinic, we would mis-estimate their PCS value if the bodily pain item was not included. Our recommendation is to use the correction and use equations that result in $R^2$of 60 percent or more. This recommendation is based on the notion that missing data is almost always more biasing than the imputation.

Table 1 shows the distribution, by decile, of the R2 values from the 4096 models that are used to predict PCS and MCS for the Veterans SF-12 from all possible patterns of missing data. These values were obtained using data from the VA 1999 Large Health Survey.

Table 1. Distribution of 4096 R^2 Values by Decile for Veterans SF-12

Decile	PCS	MCS
$R^2 = 0$ [a]	0.02	0.02
$0 < R^2 <=.1$	0.05	0.02
$.1 < R^2 <=.2$	0.12	0.05
$.2 < R^2 <=.3$	0.24	0.51
$.3 < R^2 <=.4$	0.51	1.05
$.4 < R^2 <=.5$	0.61	1.46
$.5 < R^2 <=.6$	1.66	1.54
$.6 < R^2 <=.7$	6.71	9.33
$.7 < R^2 <=.8$	23.75	20.53
$.8 < R^2 <=.9$	57.28	47.12
$.9 < R^2 <=1.0$	9.03	18.36

Note. Data based on regression weights estimated using the VA 1999 Large Survey
a. R^2 was 0 for only 1 model for each score, when all 12 items were missing.

Histograms of the R^2 values are shown below, demonstrating the highly positively skewed distributions for PCS and MCS, respectively.

[2] In the Veterans SF-12, the r^2 of items with PCS and MCS scores were .94 and .95, respectively.

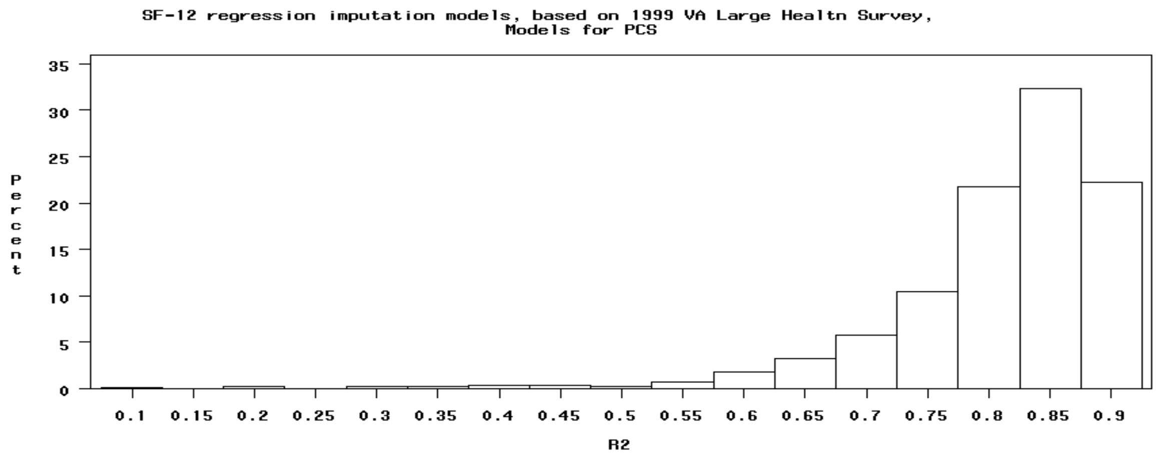

Figure 1. Histogram of R^2 values for 4096 possible SF-12 imputation models for PCS (based on 1999 VA Large Health Survey)

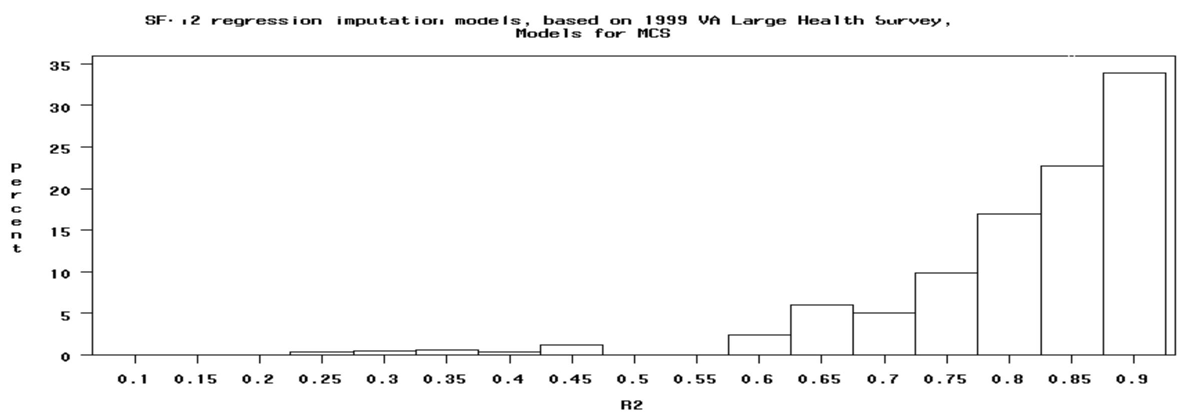

Figure 2. Histogram of R^2 values for 4096 possible SF-12 imputation models for MCS (based on 1999 VA Large Health Survey)

For PCS, the mean value of R^2 for the 4096 patterns was 0.809 (SD=.097); the median value was 0.832, and the 25th percentile was .781. For MCS, the mean value of R^2 for the 4096 patterns was 0.809 (SD=.117); the median value was 0.837, and the 25th percentile was 0.768. Note that for approximately 75% of the possible patterns of missing data, the available items predict at least 75% of the variance in the PCS or MCS scores; only about 5% of the possible patterns have R^2 that explain less than 60% of the variance in PCS or MCS. These are likely to be conservative estimates, because most cases with missing data have 8 to 11 items from the SF-12, and in general models with few missing items have higher r^2.

5. Applying the MRE Approach

Using data from the VA 1999 Large Health Survey and from baseline cohorts 1, 2, and 3 of the Medicare HOS, we have applied the MRE approach to estimating SF-12 PCS and MCS scores. The VA survey used the Veterans SF-36, which includes the Veterans SF-12, so the results are directly relevant. Because the HOS used the MOS SF-36 (which includes the MOS SF-12), and the 4 revised role items differ from those in the Veterans SF-12, the results are illustrative; they serve only to indicate the potential for estimating missing observations in future surveys that would include the Veterans SF-12.

<u>Imputing the Veterans SF-12: 1999 VA Survey</u>

The 1999 VA survey (see Kazis et al., 2000 for details), was administered to a random sample of nearly 1.5 million of 3.5 million enrollees. Nearly 65% of the sample, or 877,775 persons, responded to a survey which included the Veterans SF-36, and is embedded within, the Veterans SF-12.

For purposes in this report, we examined only the 12 items used in the Veterans SF-12. Of the 877,775 respondents, 75.5% completed all 12 items; only 2.7% omitted all 12 items. The remaining 193,479 (21.8%) respondents completed 1 (0.08%) to 11 (15.3%) of the items. Applying the MRE approach, as implemented in the SAS macro included in Appendix C, we examined the ability to recover PCS and MCS scores for those respondents with partial missing data.

Using the SAS macro described below[3], we computed PCS and MCS scores for the sample. Of the 877,775 respondents, most completed all 12 items. However, among the 193,479 cases with 1 or more missing items, we were able to compute PCS or MCS scores for 99.6% using this macro. For 98%, we were able to compute both PCS and MCS scores. Due to slight differences in the ability of the same pattern of missing items to attain comparable R^2 for PCS and MCS, there were 2691 cases for whom we could compute a PCS but not an MCS score, and 415 cases for whom we could compute an MCS but not a PCS score.

As noted, the data used in the 1999 VA survey were obtained from the Veterans SF-36. Because 93% of this sample had PCS and MCS scores from the Veterans SF-36, we were able to compare them to the imputed PCS and MCS scores from the Veterans SF-12.

[3] We used .6 as the minimum R^2 allowed for an imputation model (see below)

Table 2. Descriptive Statistics for PCS and MCS scores from 1999 VA survey

	PCS		MCS	
	Veterans SF-36	**Veterans SF-12 [a]**	**Veterans SF-36**	**Veterans SF-12 [a]**
N	824,263	862,236	824,263	859,960
Mean	35.7375	35.5583	45.1256	44.9471
SD	12.0741	12.0964	13.7648	13.7660
75th percentile	45.1334	45.0422	57.2435	57.2562
Median	34.6954	34.4487	46.0922	45.5789
25th percentile	26.0674	25.6046	34.6985	34.3890
IQR	19.0660	19.4376	22.5550	22.8672
Correlation (PCS, MCS)	.2501	.2975		

a. Imputed values were adjusted for r^2, allowing a minimum of 0.60.

Note that with the MRE imputation approach on the Veterans SF12, we were able to obtain PCS scores for an additional 37,973 respondents, and MCS scores for an additional 35,697, compared to using the half-scale rule on the SF-36.

For those respondents with scores on PCS from both SF-36 and SF-12, the mean difference in mean scores was 0.0213 (SE = .00356); for MCS, the mean difference was 0.0432 (SE = 0.00361). The correlation between PCS scores on the SF-36 and the SF-12 was 0.9643, and between MCS scores was 0.9716.

Imputing the Veterans SF-12: Medicare's Health Outcomes Survey

As noted above, because the HOS includes the MOS version of the SF-12 rather than the Veterans version, we were unable to estimate accurate PCS and MCS scores, due to differences in the role items. However, by considering the 12 items of the SF-12 that are included in the HOS SF-36, we are able to examine the patterns of missing data and determine how many cases could be imputed, granted the assumption that a missing role item would have occurred and ignoring differences due to the 2-point vs. 5-point response formats.

For this analysis, we combined all observations available to us from HOS baseline cohorts 1, 2, and 3, for a total of 879,202 persons. This number may include duplicate observations across cohorts, and does include incomplete surveys and inconsistent respondents.

Of the 879,202 respondents, 506,855 (57.6%) completed all 12 items of the SF-12. Using standard methods for scoring the SF-12, which require completion of all items, nearly half (42.4%) of HOS respondents would not have scores on PCS or MCS. Note that the majority (75%) of those respondents were missing all 12 items. By comparison, 551,877(63%) of HOS respondents provided PCS and MCS scores on the MOS SF-36, which is included in the HOS, and which can be scored using the half-scale rule (i.e., scale scores can be computed if half of the items are present, although computation of PCS and MCS scores require all 8 scales).

We then excluded from consideration respondents who had inconsistent information on birth date (2.4%), gender (1.0%) , or race (3.0%); had an invalid survey (invsrv=1; 2.8%), or whose survey disposition was incomplete (we allowed M or T 10, 11, or 31; 31.5%). This eliminated 37.3% of the 879,202, leaving 551,086 respondents for further analysis.

Of these respondents, 871 (0.2%) had none of the SF-12 items; 466,945 (84.7%) had all 12 items, and 83,270 (15.1%) had 1 to 11 items. Using the MRE approach as implemented in the SAS program in Appendix C, we would be able to impute PCS scores for 99% of those with 1 to 11 missing items, and MCS scores for 90% of those with 1 to 11 missing items.

Purposes of Sections 6, 7, and 8

The following sections 6, 7 and 8 provide the background for the theoretical foundation for the estimates used to compute missing values based upon the case-wise deletion, half scoring rule, MDE and MRE approaches. Finally the theory methods and results of the validation for the SF-12 are presented.

6. Theory and Methods for Estimates

Existing Approaches to Missing Data

In analyses involving missing item responses, where items are used to compute scale scores, there are a number of approaches that can be used. We review below three such approaches, and then propose a new approach, based on regression analysis.

Casewise deletion. The most convenient solution to missing data is simply to delete it. This solution, often referred to as casewise deletion, is a popular default in some statistical software. The result of any arithmetic operation is missing if any component is missing. Thus, when one or more items used to define a scale are missing for a case, the scale score is not computed for that case.

The problem with casewise deletion is that many observations may be lost even though there are only slight amounts of missing data. For example, a case would be lost if even 1 of the 12 items on the SF-12 were missing. A large fraction of potential cases can be eliminated in this way. For example, in the first 3 HOS baseline cohorts, about 15% of the cases with some SF-12 items had fewer than 12. In the VA 1999 Large Health Survey, about 21% of cases with any SF-12 item had fewer than 12.

The loss of so many observations raises questions about both the bias and the precision of estimates drawn from the complete cases.

Half-scoring rule. The second method of handling missing data comes from the original SF-36 reference (Ware, Kosinski, & Keller, 1993) and has a long history of use (Ware, 1976). Under the half-scoring rule, a scale is considered to be scorable if half or more of the items are present. The remaining items are for the most part prorated (i.e., assigned the mean of the items present). The PCS and MCS scores are considered scorable only if all 8 of the scales can be scored (Ware, Kosinski, & Keller, 1994).

One major limitation with the half-scoring rule is that in many cases have scales that can be scored usefully with much less data than half. Another limitation is that the method does not take into account which items are missing. If the items have varying degrees of difficulty (in the Guttman scaling sense), it does not matter if the "easiest" or the "hardest" item is missing, the rule is the same. With regard to scoring the summary scores, the rule is also conservative. Not all items are really needed for computing PCS and MCS, particularly if a relatively unimportant item is missing.

Missing Data Estimates (MDE). This method of imputation is based on extensions to Item Response Theory (e.g., Embretson & Riese, 2000) for dealing with multivariate concepts. At least 3 such extensions exist, but at this time details are unavailable. These approaches have great promise; however, they are proprietary and the documentation on them is limited (cf. Kosinski et al., 2000; QualityMetric™ at www.qualitymetric.com). Because of the current proprietary nature of the MDE approach, we do not consider it further.

New Approach to Missing Data. Here we propose a new approach to missing data, based on regression estimation. We have previously applied this method to the Veterans SF-36 for estimating scores in the context of missing data (Rogers, Qian, & Kazis, 2003). This approach is a simple modification of the approach used to construct the SF-12 from the SF-36, which involves defining an indicator variable for each response level of each SF-12 item (setting one aside, as in dummy variable methods). We propose a simple regression estimation (RE) approach, and then propose a modified version (MRE).

Regression estimates (RE). This approach is based on breaking each item down into a set of indicator (dummy) variables for the various response choices and then regressing PCS and MCS scores on these indicator variables for available items. For the Veterans SF-12, with the 5-point response scales for the 2 RP and 2 RE items, there are 47 such dummy variables. For example, the PF01 item has three responses (1=limited a lot, 2=limited a little, 3=not limited at all). Indicators are defined to indicate whether a respondent provided response 2 (pf01r2) or response 3 (pf01r3). If the respondent chooses 2, then pf01r2=1 and pf01r3=0. One indicator in each set is always omitted; we have chosen to omit the lowest response, 1.

The method then uses all available data to estimate a regression equation predicting PCS (or MCS) using only those items that are present. The following gives the complete equation assuming all items are present.

$$PCS = a + b_1 * pf02r2 + b_2 * pf02r3 + b_3 * pf04r2 + ... + b_{47} * mh4r6$$

The SF-12 is one such regression estimate based on the assumption that only 12 items are administered. Regression estimates depend on a “training” data set (which is used to obtain the weights for predicting PCS and MCS from item responses) so they are data-dependent, similar to the MDE. For the original SF-12, the training data came from the 1990 NORC survey; for the SF-12 version 2.0, data came from the 1998 NORC survey. Other subsets have also been fielded in various studies (i.e., we have used data from the VA 1999 Large Health Survey for the Veterans SF-12).

To obtain Veterans SF-12 PCS and MCS scores for cases with complete data, a regression is run where the 12 items are used to define 47 response indicators, and the response indicators are weighted using previously established regression weights from the VA 1999 Large Health Survey of Veteran Enrollees. To generalize the approach to permit estimation of PCS and MCS for cases with missing data, the same approach is used, except that, depending on the pattern of missing item responses, some weights are set to 0 (corresponding to the missing items). We have obtained, from the 1999 VA Survey, 4096 sets of weights which correspond to all possible patterns of missing data for 12 items. These weights can be applied to the user’s data, which includes cases with missing observations, to predict PCS and MCS scores, by means of the program in Appendix C.

Modified Regression Estimate. One limitation of the RE method is that the regression estimates are pulled toward the mean of the particular training data set, depending on the number and usefulness of the items available. This creates bias if the estimates are extended to outside populations or even distinct subpopulations in the original sample. The following modification corrects for this regression-to-the mean effect:

$$Y_{\text{modified}} = (\text{average}) + (Y_{\text{regression}} - \text{average}) / R$$

where R is the square root of R-squared (percent variance explained) in the regression model used and average is the average value in the training dataset. The benefits of doing this are discussed below (Section 7).

7. Theory and Methods for Validation

Further discussion of the benefits of "imputing" scores for missing data depends on two error concepts--*bias and variation*. *Bias* occurs because the estimate used differs systematically from what we would have obtained with complete data. *Variation* occurs because an estimate varies around the expected answer, due to sampling. Theoretically, it helps to conceptualize what the answer would have been if there were an infinite number of observations with the same missing data phenomena that are seen in the finite data.

Error = bias + variation = (infinite answer - true answer) + (sample answer - infinite answer)

As the sample size increases, the first term remains the same, but the last term approaches zero according to the law of large numbers. Accordingly, bias is much more of a threat in large samples, but variation is more of a threat in small samples. In large samples, we need to take care with imputation or case exclusion because of the dangers of drawing an incorrect conclusion with a false sense of precision. In small samples we need to be concerned with the unnecessary deletion of observations. Whether the sample is large or small depends on both the study and what is being compared. In the case of the HOS, the sample is very large if we are following the health of patients in HMOs generally, but smaller if we are comparing health plans.

In a given situation, bias and variance arise because of different aspects of analysis, so we can create a formal trade-off and attempt to minimize a combination of the two. The combination usually encountered is Mean Squared Error (MSE) which is defined

$$\text{MSE} = \text{bias}^2 + \text{variation}^2$$

To give this problem more analytic structure, we have two options for each missing data strategy--we can include the observations with their missing data estimates, or we can exclude them. In addition, we can weight them. A weight of 0 corresponds to excluding them, and a weight of 1.0 is equivalent to including them. Given N1 samples with complete cases and N2 samples which could be imputed with squared *bias h* and variations with *variance v*:

$$\text{Bias contribution} = (N2*W/(N1+W*N2))^2\ h$$

$$\text{Variation contribution} = (N1+W^2\ N2)/(N1+W\ N2)^2\ v,$$

where N1= samples with complete cases, N2 = samples with incomplete cases, W = weight, h = squared bias, and v = squared standard deviation.

This assumes that the variation contributes about the same amount for complete as well as incomplete cases. Unless the amount of missing data is extreme, the variation of the imputed observations is about the same as the complete cases. In addition, it is helpful to express both h and v in standard terms--the only thing that really matters is the ratio h/v and the sample sizes N1 and N2.

Illustration with systematic planned missing data. To illustrate these terms with practical data, imagine that we simulate the planned omission of the pf01 item under the half-scale rule and we evaluate Veterans SF-36 PCS for a population mean, and a comparison of health plan baseline scores.

For the population mean, the bias in half-scale Veterans SF-36 PCS was $B_K = 0.6079$ points on average and the standard deviation of (half-scale Veterans SF-36 PCS - true value) was $SD_K = 0.550$. The standard deviation of Veterans SF-36 PCS was 11.74. The ratio h/v is $.6079^2/11.74^2$ or 0.00268. Suppose that N1 is 650 and N2 is 350. Then W should be about 0.5 and imputation is better than no imputation. However, if N1=6500 and N2=3500 we are much better off not imputing.

For comparison of health plan baseline scores, we run an analysis of variance (ANOVA) of the difference (PCS with half-scale scoring - PCS gold standard or complete data) on health plan baseline ID. We get a SS(plan) of 1036.77 with F = 10.83, so

$$h/v = (\text{SS of effect}) * (F-1)/F / (N * SD^2)$$

$$h/v = 1036.77 * ((10.83-1)/10.83)/ (289650 * 11.74^2) = 0.0000236$$

For health plan sizes of about 1000, as in the HOS, the optimal value of W is very close to 1, which suggests that we should impute and use the observations. This does not mean that half-scale imputing is better than other types of imputing (see results), but it does mean that if we are missing PF01 and we are given a choice of case-wise deletion or using half-scale, we should use half-scale imputation.

Problems with various imputation methods can be traced mostly to the fact that items have unique content as well as error. For example, within the PF scale of the SF-36, PF8 is an item that describes walking several blocks and PF9 describes limitations in walking one block. Both are part of the physical functioning scale. The Pearson correlation (same as Spearman) between the two items at time 1 is 0.58. Comparing the two waves, the two PF9 items are correlated 0.61, and the two PF8 items 0.65, but the cross correlations are 0.54 and 0.61. Cross correlations are only slightly lower, suggesting that just over 90% of the variance is shared (.58/.63) and a little less than 10% is unique. For PF6 (bending, kneeling, and stooping) in relation to PF9, a similar technique tells us that two-thirds of the variance is overlapping and one-third is unique.

8. Validation Results for PCS and MCS from the Veterans SF-12

Methods used to validate the Veterans SF-36 in the 1999 VA survey have been described previously (Rogers, Qian, & Kazis, 2003); a portion of these results which pertain to the SF-12 is shown here. Table 3 describes the observations which were useable under various scenarios. About 2/3rd's of observations were usable under the casewise deletion approach. Using the half-scale rule for the SF-36 resulted in an improvement to 95% usable cases; with the MRE approach (allowing an r^2 of .5 or larger), nearly all cases were usable.

Table 3: Usable Observations under Various Scenarios

Method	**Number Scorable**	**% of possible cases**
Casewise deletion	587,642	68.04
Half-scale	824,301	95.44
MRE (PCS, r^2>0.5)	863,565	99.99
MRE (MCS, r^2>0.5)	861,704	99.78

In Tables 4a and 4b, 'SF12' is the classical SF-12; 'SF1' is the single gh1 item of overall health from excellent to poor (which cannot be scored with the MDE estimator).

*Table 4A: PCS bias properties (h/v) of Regression Estimates (x1000).**

		VISNS		Health Conditions		Demographics	
Values	**R-sq**	**RE**	**MRE**	**RE**	**MRE**	**RE**	**MRE**
SF12	93.8	0.1215	0.0623	1.7225	0.9292	0.1720	0.0890
SF1*	51.9	7.3217	1.3029	110.1200	38.3420	22.6710	10.5930

*Table 4B: MCS bias properties (h/v) of Regression Estimates (x1000).**

		VISNS		Health Conditions		Demographics	
Values	**R-sq**	**RE**	**MRE**	**RE**	**MRE**	**RE**	**MRE**
SF12	95.2	0.0191	0.0111	1.5226	0.9417	0.5541	0.4587
SF1*	26.0	6.4660	4.2440	682.1100	154.1600	138.7700	52.0790

These tables show that the missing value estimators seem to be usable if as few as 3 items are present, as long as they draw from the three main concepts, physical, bodily pain and mental health, i.e., PF, BP, and MH. It is possible that other configurations would work, but we did not test them. With a bias property of h/v = (SS of effect) * (F-1)/F / (N * SD^2)
= 4.3 x 10^{-3}, the SF-1 (the GH1 item of the SF-36) would have a typical *error* of 0.76 points as an estimate of PCS, about a third of the health plan PCS effect (determined to be 1.65 points by variance components). This means that about 21% of the variation is "off concept" relative to the PCS, but 79% is on-concept in this extreme case. For determination of disease means, the error is 1.77 points, but these means often differ by 5-10 points.

The advantage of the MRE compared to the simple regression estimator becomes more important when we are dealing with more extreme imputation of missing values, particularly for MCS. ‘SF1’ would lead to errors of several points, but ‘SF3a’ seems to be quite usable with errors of about 0.5 points, typically.

We can't say much about the MDE in these analyses because we did not have access to a convenient algorithm to score it for simulation. We can however, compare the behavior of the MDE and MRE in naturally missing data. These indicate that the two estimators are fairly close, differing by a mean of -0.012 (MDE is lower) with a SD of 0.40 between them. That suggests they will be within 1 point of each other almost all the time. For example, when ‘PF01’ is missing, MDE is lower by 0.13 points, and if ‘PF10’ is missing, MDE is higher by .067 points. *A multiple regression of the MDE on the half-scale rule and the MRE suggests that the MDE is closer to the half-scale rule than it is to the MRE, particularly for MCS. However, the correlation between the half-scale rule, the MDE, and the MRE gives coefficients of 0.9997 and higher.*

Tables 5A and 5B indicate that the PCS and MCS bias due to naturally occurring data compared with the MRE as the standard is superior to the half scale rule for VISNs, conditions and demographics.

Table 5A : PCS Bias due to naturally missing data compared with the MRE approach as the standard

		PCS Bias (h/v) x 1000		
Imputation Algorithm	**MeanBias**	**VISNS**	**Conditions**	**Demographics**
Complete case	-3.17	0.1526	1.4263	2.0598
Half-Scale rule	-4.28	0.0946	0.4201	1.5154

Table 5B: MCS Bias due to naturally missing data compared with the MRE approach as the standard

		PCS Bias (h/v) x 1000		
Imputation Algorithm	**MeanBias**	**VISNS**	**Conditions**	**Demographics**
Complete case	-2.55	0.1350	0.1253	0.3098
Half-Scale rule	-2.93	0.0884	0.1175	0.6197

The “MeanBias” column in tables 5A and 5B describes how cases that cannot be scored with the imputation algorithm differ from those that can be scored. The impact is proportional to this number times the percentage that is missing. For table 5A that describes PCS, the half-scale rule gives about 92% of the cases for PCS, so the mean bias associated with not scoring it is 8% times -3.26 or about -0.26 bias points. For PCS the equivalent error in points is 0.15. For MDE, the MDE gives 96.6% cases, so the mean bias associated with not scoring is 3.4% times -2.42 or about -0.08 bias points which is equivalent to about the same PCS error in points. For health plan comparisons the MDE approach gives about 0.15 point error for PCS compared with the MRE approach. For table 5B, the mean bias associated with not scoring using the half scale rule is 8% times -3.49 or about -0.28 bias points for MCS which is equivalent to about 0.18 point error. For health plan comparisons for the MDE approach, there is an error of 0.25 points for MCS compared with the MRE approach.

The remaining columns should be interpreted similar to the systematic planned tables above. That is, the biases shown have been multiplied by 1000. Although none of these biases is serious, they offset typical biases from imputation. They also suggest that the MRE approach is less biased compared to the half scale rule and the MDE, although marginally for the MDE.

9. Implications for Analysis

Based on the above analysis, the MRE is our preferred method if the goal is to replicate original values of the SF-36 summaries in a point in time. We should state that the half-scale rule is not adequate for the purposes of imputation. The MDE while almost comparable to the MRE is not fully available to us since it requires enormous computer resource requirements for scoring and unknown algorithms that are not available to the public. We conclude that the MRE method is the more reasonable approach for estimating individual scale values (e.g. PF, RP, etc) of the MOS and Veterans SF-36 and the SF12.

Given that we have selected a preferred method and know about the bias typically associated with it, how should estimation be done using this approach? The following points should be kept in mind:

a. For complete cases, use the complete case value.

b. For incomplete cases, use the MRE method, so long as the MRE reaches the threshold of acceptable performance--we suggest an R-squared of 0.6 or greater.

Because the MRE approach results in very little bias--even when we have used fairly extreme cases of missing values. We do not suggest weighting the imputed data. The observations imputed with the MRE should be used without weights..

10. Some conclusions about Imputation Approaches

In recent years other more sophisticated approaches have been developed for imputation of the SF-12. The MDE is rooted in a sound and currently popular theory of scale psychometrics (Item Response Theory). Its main disadvantages are the complicated and proprietary nature of the software. The regression imputation is based on older regression technologies, but is an order of magnitude more complicated than the half-scale rule. The MDE requires a complicated software program to run effectively, and the means to do that within popular computer software has evolved with the speed of the computers and the sophistication of software programs (e.g. SAS and STATA) The MRE, described here, employs a simple yet effective correction for regression to the mean that makes the regression estimate more general (and therefore less biased) than it would otherwise be.

We found that failing to impute resulted in more bias in the results than imputing the results. The MRE has relatively small imputation biases which cancel out in naturally missing data, but the biases due to not imputing (and losing the cases) are consistent. The more that data are imputed, the less biased the overall answers will be, and they will also be more accurate due to the additional sample size. Although this statement applies most to the MRE, it would also apply to the MDE when compared to the half-scale rule.

Our ability to directly compare the MRE and the MDE was limited given that the algorithm was not available to us. The evidence though in the results available to us using the Web site that calculates results using the MDE approach through QualityMetric™ suggests the MRE methodology is better. The MRE method imputes more cases and so should be both less biased and lower in variance. In addition, the correlation analysis produced better agreement between the MRE and follow-up data than between the MDE and follow-up data, even if the MDE was used for follow-up. The MDE appears to retain some affinity to the half-scale rule--though it is far better than the half-scale rule. We do not take the use of Item Response Theory (IRT) to be an advantage of the MDE, but we do not know it is a disadvantage either. The MDE is just another approach using an IRT statistical model that needs to be trained. These negatives could be offset by possible advantages of the MDE in non-HOS populations, since the MDE presumably had a more diverse training set and therefore might be more generalizable.

The mean bias of unimputed cases was negative in all cases. This implies that when patients don't fill out lots of items, their health is typically poorer than when they do fill out all or most of the items. However, the illustration in methods for this report suggested that was not true for every item. Neither the MRE (nor we think the MDE) address the fundamental question of whether the naturally missing nature of the items conveys information beyond being missing at random, once the values of the other items have been properly taken into account. Nor have we addressed the interesting question of whether missing data somehow signals impending change in the SF-12.

11. Alpha Testing of the Manual, Users Guide and Computer Program

Evaluation of the Manual and Users' Guide for the Veterans SF-12 Imputation Program

We distributed the Veterans SF-12 Imputation Manual and User's Guide to six programmers with varying degrees of proficiency in SAS (from beginner to advanced). Each was given a zip file, distribution.zip, which contained the Manual, User's Guide, sample program, macro, and data files for the Veterans' SF-12 imputation program. In addition, each was asked to complete the questionnaire shown in Appendix E.

Results are shown in attachment labeled Table 6 (Excel attachment labeled "Table 6 Results of Evaluation"). The majority of users found the manual good to very good (Question 1) ; most felt it was complete (Question 1a). They agreed that the theory and methods were adequately explained (Question 1e) and that the scoring methodology for the Veterans SF-12 was explained well (Question 1f).

Users felt that the User's Guide was very good (Question 2), clearly written (Question 2a), and was helpful in running the program (Question 2b). Most users were able to run the program in about under one hour, although one required three hours. All felt that the program was fairly easy to run.

Appendix A. Weights for Scoring Veterans SF-12 PCS and MCS

Veterans SF-12 ITEM	Variable Label	RESPONSE CHOICE	LABEL	PCS COEFF	MCS COEFF
Constants				47.226630	44.856200
Moderate activities	**PF02**	Limited a lot	--	--	--
(Physical Functioning)		Limited a little	pf2r2	3.209097	-1.741941
		Not limited at all	pf2r3	6.440926	-3.391449
Climbing several flights	**PF04**	Limited a lot	--	--	--
of stairs		Limited a little	pf4r2	3.841436	-1.893174
(Physical Functioning)		Not limited at all	pf4r3	6.875059	-3.358263
Accomplished less than you	**VRP2**	None of the time	--	--	--
would like		A little of the time	vrp2r2	-2.295770	0.770424
(Role Limitations because of		Some of the time	vrp2r3	-4.220704	1.342969
Physical Problems)		Most of the time	vrp2r4	-5.869204	1.843018
		All of the time	vrp2r5	-6.451106	2.113603
Limited in the kind of work	**VPR3**	None of the time	--	--	--
or activities		A little of the time	vrp3r2	-2.853384	0.898016
(Role Limitations because of		Some of the time	vrp3r3	-4.751619	1.519380
Physical Problems)		Most of the time	vrp3r4	-6.292369	1.932001
		All of the time	vrp3r5	-6.834621	2.089988
How much pain interferes	**BP2**	Not at all	--	--	--
with normal work		A little bit	bp2r2	-3.767011	0.724378
(Pain)		Moderately	bp2r3	-6.888286	1.289420
		Quite a bit	bp2r4	-9.701818	1.752278
		Extremely	bp2r5	-12.553300	2.261750
In general, you would say	**GH1**	Excellent	--	--	--
your health is		Very good	gh1r2	-1.422927	0.006179
(General Health)		Good	gh1r3	-3.200699	-0.032633
		Fair	gh1r4	-5.668607	-0.151991
		Poor	gh1r5	-7.623203	-0.410722
Have a lot of energy	**VT2**	All of the time	--	--	--
(Vitality)		Most of the time	vt2r2	-0.487705	-0.863361
		A good bit of the time	vt2r3	-1.054558	-1.997290
		Some of the time	vt2r4	-1.570157	-3.313938
		A little of the time	vt2r5	-2.004446	-4.671423
		None of the time	vt2r6	-2.565244	-6.016106
How much time health	**SF2**	All of the time	--	--	--
interferes w/social activities		Most of the time	sf2r2	0.214456	2.148606
(Social Functioning)		Some of the time	sf2r3	0.270629	4.989030
		A little of the time	sf2r4	0.523565	7.583853
		None of the time	sf2r5	0.772322	10.251920

Veterans SF-12 ITEM	Variable Label	RESPONSE CHOICE	LABEL	PCS COEFF	MCS COEFF
Constants				47.226630	44.856200
Accomplished less than you would like (Role Limitations because of Emotional Problems)	**VRE2**	All of the time	--	--	--
		Most of the time	vre2r2	1.863268	-3.867584
		Some of the time	vre2r3	3.491722	-7.704990
		A little of the time	vre2r4	4.604420	-10.290840
		None of the time	vre2r5	4.502007	-10.038810
Didn't do work or other activities as carefully as usual (Role Limitations because of Emotional Problems)	**VRE3**	All of the time	--	--	--
		Most of the time	vre3r2	1.213867	-3.052609
		Some of the time	vre3r3	2.227551	-5.676195
		A little of the time	vre3r4	2.839852	-7.568439
		None of the time	vre3r5	2.273264	-6.684413
Felt calm and peaceful (Mental Health)	**MH3**	None of the time	--	--	--
		A little of the time	mh3r2	0.509143	-1.945028
		Some of the time	mh3r3	1.250000	-3.920049
		A good bit of the time	mh3r4	2.136413	-6.051385
		Most of the time	mh3r5	3.068895	-8.191803
		All of the time	mh3r6	3.758398	-9.805100
Felt downhearted and blue (Mental Health)	**MH4**	None of the time	--	--	--
		A little of the time	mh4r2	-0.733526	2.825623
		Some of the time	mh4r3	-1.840210	6.163902
		A good bit of the time	mh4r4	-3.020777	9.500628
		Most of the time	mh4r5	-3.943621	12.128690
		All of the time	mh4r6	-4.854536	14.706530

Appendix B. Use of the SAS© Macro to Impute PCS & MCS for the Veterans SF-12

To use the SAS macro to impute PCS and MCS scores for the Veterans SF-12, use the following steps. They are illustrated with a sample program below (Appendix B.1), for which the SAS log (Appendix B.2) and list (Appendix B.3) files are then shown.

The data used in the example are included, as SAS system file sample12.sas7bdat. These data are a 1% random sample (n=8,637) of cases with at least 1 Veterans SF-12 item, extracted from the 1999 VA survey, which used the Veterans (not the MOS) SF-36. However, the variable names for the SF-12 items have been revised to reflect those used by the Health Outcomes Survey. In addition to the SF-12 items, the sample data include age, gender, and PCS and MCS scores from the Veterans SF-36. Of the 8,637 respondents, only 4.6% were women. Their mean age was 63 (SD = 13.6, range 20 to 97). Most (77.7%) completed all SF-12 items; 15.7% completed 11 of 12.

0. Create a SAS program that reads your HOS data, with formats, etc.

1. Include the imputation macro, e.g.,

```
%include ‘LOCATION1\v-sf12-impute1.2.sas’;
```

where LOCATION1 is a pathname indicating where the imputation program is located.

2. Specify the library name where the PCS and MCS weights are stored, e.g.,

```
Libname X ‘LOCATION2’;
```

The above statement assigns ‘X’ to the pathname specified by LOCATION2. Note that LOCATION 1 in Statement 1 and LOCATION2 in this statement can differ.

3. Include a statement in your SAS program to execute the imputation macro,

```
%vsf12imp(
 indata= <name of your SAS dataset containing SF12 items>,
 idvar= <name of a case identification variable [default=id]>,
 minr2= <minimum value of R^2 for imputation [default=.6]>,
 PCS_WTS = <X.PCS, if X is libname assigned in Step 2 above>,
 MCS_WTS = <X.MCS, if X is libname assigned in Step 2 above>,
 Validity= <0=no validity check [default], 1=validity check>,
 Outdata= <SAS name for output dataset [default= _imputed]>
 );
```

Example: `%vsf12imp(indata=mydata, PCS_WTS=X.PCS, MCS_WTS=X.MCS);`

4. Submit the SAS program.

5. The results of the imputation program, output in the dataset defined by the “outdata” parameter in the macro execution statement (Step 3) can be saved or merged with other data for purposes of analysis.

Appendix B1. Sample SAS program

```
options NOcenter ;
title '\chqoer\SF12\sample             20 sep 04';
title2 'CMS/HOS, testing the imputation on Sample Data with Veterans SF-12';

libname WT 'c:\RS\SF12';            /* location of PCS & MCS weights */
Libname X  'C:\RS\sf12';            /* Location of input data */
%let TST = X.SAMPLE12;                    /* Name of input data */
%include  'c:\RS\sf12\v-sf12-impute1.2.sas';    /* name/path of SAS Imputation
macro */

proc format;
  value SEXF 1='Male' 2='Female';

* ------------------------------------------------------- *;
* Input test data
* ------------------------------------------------------- *;
 /* NOTE: Rename of HOS Cohort 3 variables to SPECIFIED Veterans SF12 names */
data test ;
 set &TST (rename=(C3modact=PF02 C3clmbsv=PF04 C3pacmpl=VRP2 C3plmtkw=VRP3
             C3pnintf=BP2 C3genhth=GH1 C3energy=VT2 C3sclact=SF2
             C3Eacmpl=VRE2 C3entcrf=VRE3 C3pceful=MH3 C3blsad=MH4 ));

  /* NOTE: dataset input to imputation macro must be sorted by user-defined
IDVAR variable */
proc sort; by ID;

* ------------------------------------------------------- *;
*  EXECUTE THE IMPUTATION MACRO
* ------------------------------------------------------- *;

  %VSF12IMP(indata= test, idvar= id, pcs_wts=WT.PCS, mcs_wts=WT.MCS,
           validity=1, omit=1,
           outdata= X._testimp);

* ------------------------------------------------------- *;
* Merge imputed SF-12 scores with Original data
* ------------------------------------------------------- *;
data work;
 merge test X._testimp ;
  by id ;

 title 'Merge of Original data and imputed Veterans SF-12';

 *** TESTING: differences between various scores *** ;
   * SF36 scores vs. unadjusted imputed scores;
  d_pcs2 = C3PCS - pcs12 ;
  d_mcs2 = C3MCS - mcs12 ;
   * SF36 scores vs. adjusted imputed scores;
  d_pcs3 = C3PCS - pcs12_adj ;
  d_mcs3 = C3MCS - mcs12_adj ;
  label
    d_pcs2 = 'PCS(SF36) - PCS12' d_mcs2 = 'MCS(SF36) - MCS12'
    d_pcs3 = 'PCS(SF36) - PCS12_adj' d_mcs3 = 'MCS(SF36) - MCS12_adj';
```

```
54
55   * ------------------------------------------------------- *;
56   proc means n mean std min max;
57
58   proc means n mean stderr t prt ;
59     title3 'Differences among estimated scores';
60     var d_pcs2 d_pcs3   d_mcs2 d_mcs3;
61
62   proc corr;
63     title3 'Correlations among all scores';
64     var C3pcs pcs12 pcs12_adj C3mcs mcs12 mcs12_adj ;
65
66   * ------------------------------------------------------- *;
67   run;
```

Comments on the Sample Program

Line numbers	Comments
1-3	SAS options, program titles
5-9	Identify locations and names of input dataset, imputation macro, and PCS/MCS weight files
11-12	Formats for variables in input data
18-21	Create a dataset reading the sample data, and rename the HOS variable names for SF-12 items to the names required by the imputation macro
25	Sort the dataset by the case identifier (here, ID), as required by the imputation macro
31-33	Execute the imputation macro. For further information on defining the required macro variables, see Appendix C (lines 20 – 53)
38-40	Merge the imputed PCS & MCS scores with the sample data, using the case identifier.
44-53	Define and label variables computing the discrepancy between Veterans SF-36 PCS and MCS scores and the Veterans SF-12 scores, with the imputation.
56-64	Compute means, t-tests, and correlations among SF-36 and SF-12 scores.

Note. The data used in this sample program and included with this documentation are extracted from the 1999 VA survey, which used the Veterans SF-36. The HOS currently uses the MOS SF-36, and the 2-point role items instead of the 5-point role items of the Veterans SF-36. The items of the SF-12 were assigned the names used in HOS data, cohort 3 baseline.

Appendix B2. SAS Log file from sample program

Notes: Line numbers are created by SAS.

```
598  title '\chqoer\SF12\sample             20 sep 04';
599  title2 'CMS/HOS, testing the imputation on Sample Data with Veterans SF-12';
600
601   Libname X  'C:\RS\sf12';
NOTE: Libname X refers to the same physical library as WT.
NOTE: Libref X was successfully assigned as follows:
      Engine:        V8
      Physical Name: C:\RS\sf12
601!                                                 /* Location of input data */
602   %let TST = X.SAMPLE12;                         /* Name of input data */
603   libname WT 'c:\RS\SF12';
NOTE: Libname WT refers to the same physical library as X.
NOTE: Libref WT was successfully assigned as follows:
      Engine:        V8
      Physical Name: C:\RS\sf12
603!                                                 /* location of PCS & MCS weights */
604   %include  'c:\RS\sf12\v-sf12-impute1.2.sas';    /* path of SAS Imputation macro */
1136
1137  proc format;
1138    value SEXF 1='Male' 2='Female';
NOTE: Format SEXF is already on the library.
NOTE: Format SEXF has been output.
1139
1140  * ------------------------------------------------------- *;
1141  * Input test data
1142  * ------------------------------------------------------- *;
1143   /* NOTE: Rename of HOS Cohort 3 variables to SPECIFIED Veterans SF12 names */

NOTE: PROCEDURE FORMAT used:
      real time           0.01 seconds
      cpu time            0.01 seconds


1144  data test ;
1145   set &TST (rename=(C3modact=PF02 C3clmbsv=PF04
1146               C3pacmpl=VRP2 C3plmtkw=VRP3 C3pnintf=BP2 C3genhth=GH1 C3energy=VT2
C3sclact=SF2
1147               C3Eacmpl=VRE2 C3entcrf=VRE3 C3pceful=MH3 C3blsad=MH4 ));
1148
1149    /* NOTE: dataset input to imputation macro must be sorted by user-defined IDVAR
variable */

NOTE: There were 8637 observations read from the dataset X.SAMPLE12.
NOTE: The data set WORK.TEST has 8637 observations and 17 variables.
NOTE: DATA statement used:
      real time           0.01 seconds
      cpu time            0.01 seconds


1150  proc sort; by ID;
1151
1152  * ------------------------------------------------------- *;
```

```
1153  *  EXECUTE THE IMPUTATION MACRO
1154  * ------------------------------------------------------- *;
1155
1156    %VSF12IMP(indata= test, idvar= id, pcs_wts=WT.PCS, mcs_wts=WT.MCS,
1157             validity=1, omit=1,
1158             outdata= X._testimp);
NOTE: There were 8637 observations read from the dataset WORK.TEST.
NOTE: The data set WORK.TEST has 8637 observations and 17 variables.
NOTE: PROCEDURE SORT used:
      real time           0.01 seconds
      cpu time            0.01 seconds





*****************************************************************
          Veterans SF-12 Imputation Program for HOS
            Health Outcomes Technologies Program
            Boston University School of Public Health
            Program Version 1.1, September 2004

         Supported by NCQA/CMS, Boston University, and
 the Research Services of the US Department of Veterans Affairs

              Name of dataset for analysis:  test
              Case identifier:               id
              Minimum R2 for imputation:     .6

              PCS weights are read from:     WT.PCS
              MCS weights are read from:     WT.MCS

              Validity check is: ON

              Cases with all SF-12 items missing are:  DELETED

*****************************************************************
NOTE: DATA statement used:
      real time           0.00 seconds
      cpu time            0.00 seconds


NOTE: There were 8637 observations read from the dataset WORK.TEST.
NOTE: The data set WORK._SF12SCAL has 8637 observations and 60 variables.
NOTE: DATA statement used:
      real time           0.06 seconds
      cpu time            0.06 seconds


NOTE: There were 8637 observations read from the dataset WORK._SF12SCAL.
NOTE: The data set WORK._NE1 has 8637 observations and 49 variables.
NOTE: DATA statement used:
      real time           0.10 seconds
      cpu time            0.10 seconds
```

```
NOTE: There were 8637 observations read from the dataset WORK._NE1.
NOTE: The data set WORK._NE1 has 8637 observations and 49 variables.
NOTE: PROCEDURE SORT used:
      real time           0.04 seconds
      cpu time            0.04 seconds


NOTE: There were 4096 observations read from the dataset WT.PCS.
NOTE: The data set WORK._PCSUSE has 3964 observations and 51 variables.
NOTE: DATA statement used:
      real time           0.01 seconds
      cpu time            0.01 seconds


NOTE: There were 8637 observations read from the dataset WORK._NE1.
NOTE: There were 3964 observations read from the dataset WORK._PCSUSE.
NOTE: The data set WORK._PCSI has 8629 observations and 7 variables.
NOTE: DATA statement used:
      real time           0.04 seconds
      cpu time            0.04 seconds


NOTE: There were 4096 observations read from the dataset WT.MCS.
NOTE: The data set WORK._MCSUSE has 3905 observations and 51 variables.
NOTE: DATA statement used:
      real time           0.01 seconds
      cpu time            0.01 seconds


NOTE: There were 8637 observations read from the dataset WORK._NE1.
NOTE: There were 3905 observations read from the dataset WORK._MCSUSE.
NOTE: The data set WORK._MCSI has 8607 observations and 7 variables.
NOTE: DATA statement used:
      real time           0.03 seconds
      cpu time            0.03 seconds


NOTE: There were 8629 observations read from the dataset WORK._PCSI.
NOTE: The data set WORK._PCSI has 8629 observations and 7 variables.
NOTE: PROCEDURE SORT used:
      real time           0.01 seconds
      cpu time            0.01 seconds


NOTE: There were 8607 observations read from the dataset WORK._MCSI.
NOTE: The data set WORK._MCSI has 8607 observations and 7 variables.
NOTE: PROCEDURE SORT used:
      real time           0.03 seconds
      cpu time            0.03 seconds


NOTE: There were 8629 observations read from the dataset WORK._PCSI.
NOTE: There were 8607 observations read from the dataset WORK._MCSI.
NOTE: The data set X._TESTIMP has 8630 observations and 7 variables.
NOTE: DATA statement used:
      real time           0.03 seconds
      cpu time            0.03 seconds
```

```
NOTE: PROCEDURE CONTENTS used:
      real time           0.00 seconds
      cpu time            0.00 seconds


NOTE: There were 25 observations read from the dataset X._TESTIMP.
NOTE: PROCEDURE PRINT used:
      real time           0.00 seconds
      cpu time            0.00 seconds


NOTE: There were 8630 observations read from the dataset X._TESTIMP.
NOTE: PROCEDURE MEANS used:
      real time           0.01 seconds
      cpu time            0.01 seconds


NOTE: There were 8630 observations read from the dataset X._TESTIMP.
NOTE: PROCEDURE CORR used:
      real time           0.01 seconds
      cpu time            0.01 seconds


NOTE: There were 8630 observations read from the dataset X._TESTIMP.
NOTE: PROCEDURE UNIVARIATE used:
      real time           0.03 seconds
      cpu time            0.03 seconds


NOTE: There were 8630 observations read from the dataset X._TESTIMP.
NOTE: PROCEDURE FREQ used:
      real time           0.03 seconds
      cpu time            0.03 seconds


NOTE: Deleting WORK._NE1 (memtype=DATA).
NOTE: Deleting WORK._PCSUSE (memtype=DATA).
NOTE: Deleting WORK._MCSUSE (memtype=DATA).
NOTE: Deleting WORK._SF12SCAL (memtype=DATA).
NOTE: Deleting WORK._PCSI (memtype=DATA).
NOTE: Deleting WORK._MCSI (memtype=DATA).
NOTE: PROCEDURE DATASETS used:
      real time           0.06 seconds
      cpu time            0.06 seconds


NOTE: There were 8630 observations read from the dataset X._TESTIMP.
NOTE: There were 8637 observations read from the dataset WORK.TEST.
NOTE: The data set WORK._VAL has 8637 observations and 23 variables.
NOTE: DATA statement used:
      real time           0.03 seconds
      cpu time            0.03 seconds


NOTE: There were 8637 observations read from the dataset WORK._VAL.
```

```
NOTE: PROCEDURE CORR used:
      real time           0.01 seconds
      cpu time            0.01 seconds


NOTE: There were 8637 observations read from the dataset WORK._VAL.
NOTE: PROCEDURE CORR used:
      real time           0.01 seconds
      cpu time            0.01 seconds


NOTE: There were 6710 observations read from the dataset WORK._VAL.
      WHERE (IMPUTE_P=0) and (IMPUTE_M=0);
NOTE: PROCEDURE CORR used:
      real time           0.03 seconds
      cpu time            0.03 seconds


NOTE: There were 1920 observations read from the dataset WORK._VAL.
      WHERE (IMPUTE_P=1) or (IMPUTE_M=1);
NOTE: PROCEDURE CORR used:
      real time           0.03 seconds
      cpu time            0.03 seconds

NOTE: Deleting WORK._VAL (memtype=DATA).
NOTE: PROCEDURE DATASETS used:
      real time           0.01 seconds
      cpu time            0.01 seconds



      * -------------------------------------------------------- *;
      * Merge imputed SF-12 scores with Original data
      * -------------------------------------------------------- *;



*****************************************************************
       --> End of Veterans SF-12 Imputation Program <--
*****************************************************************
NOTE: DATA statement used:
      real time           0.00 seconds
      cpu time            0.00 seconds


      data work;
       merge test X._testimp ;
        by id ;

       title 'Merge of Original data and imputed Veterans SF-12';

       *** TESTING: differences between various scores *** ;
         * SF36 scores vs. unadjusted imputed scores;
        d_pcs2 = C3PCS - pcs12 ;
        d_mcs2 = C3MCS - mcs12 ;
         * SF36 scores vs. adjusted imputed scores;
        d_pcs3 = C3PCS - pcs12_adj ;
```

```
1175    d_mcs3 = C3MCS - mcs12_adj ;
1176    label
1177      d_pcs2 = 'PCS(SF36) - PCS12' d_mcs2 = 'MCS(SF36) - MCS12'
1178      d_pcs3 = 'PCS(SF36) - PCS12_adj' d_mcs3 = 'MCS(SF36) - MCS12_adj';
1179
1180  * ------------------------------------------------------- *;

NOTE: Missing values were generated as a result of performing an operation on missing
values.
      Each place is given by: (Number of times) at (Line):(Column).
      371 at 1171:18   371 at 1172:18   371 at 1174:18   371 at 1175:18
NOTE: There were 8637 observations read from the dataset WORK.TEST.
NOTE: There were 8630 observations read from the dataset X._TESTIMP.
NOTE: The data set WORK.WORK has 8637 observations and 27 variables.
NOTE: DATA statement used:
      real time           0.03 seconds
      cpu time            0.03 seconds


1181  proc means n mean std min max;
1182

NOTE: There were 8637 observations read from the dataset WORK.WORK.
NOTE: PROCEDURE MEANS used:
      real time           0.04 seconds
      cpu time            0.04 seconds


1183  proc means n mean stderr t prt ;
1184    title3 'Differences among estimated scores';
1185    var d_pcs2 d_pcs3   d_mcs2 d_mcs3;
1186

NOTE: There were 8637 observations read from the dataset WORK.WORK.
NOTE: PROCEDURE MEANS used:
      real time           0.01 seconds
      cpu time            0.01 seconds


1187  proc corr;
1188    title3 'Correlations among all scores';
1189    var C3pcs pcs12 pcs12_adj C3mcs mcs12 mcs12_adj ;
1190
1191  * ------------------------------------------------------- *;
1192  run;

NOTE: There were 8637 observations read from the dataset WORK.WORK.
NOTE: PROCEDURE CORR used:
      real time           0.01 seconds
      cpu time            0.01 seconds
```

Appendix B3. SAS List file from sample program

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                 18
***  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***       17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

*   X._testimp contains Veterans SF-12 PCS & MCS scores (with MRE imputation)   *

The CONTENTS Procedure

Data Set Name: X._TESTIMP                                        Observations:          8630
Member Type:   DATA                                              Variables:             7
Engine:        V8                                                Indexes:               0
Created:       17:39 Monday, September 20, 2004                  Observation Length:    56
Last Modified: 17:39 Monday, September 20, 2004                  Deleted Observations: 0
Protection:                                                      Compressed:            NO
Data Set Type:                                                   Sorted:                NO
Label:

          -----Engine/Host Dependent Information-----

Data Set Page Size:          8192
Number of Data Set Pages:    60
First Data Page:             1
Max Obs per Page:            145
Obs in First Data Page:      113
Number of Data Set Repairs: 0
File Name:                   C:\RS\sf12\_testimp.sas7bdat
Release Created:             8.0000M0
Host Created:                WIN_PRO

     -----Alphabetic List of Variables and Attributes-----

#    Variable     Type    Len    Pos    Label
ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ
5    IMPUTE_M     Num       8     32    MCS imputed? (1=yes)
2    IMPUTE_P     Num       8      8    PCS imputed? (1=yes)
6    MCS12        Num       8     40    MCS (imputed)
7    MCS12_adj    Num       8     48    MCS (imputed), adjusted
3    PCS12        Num       8     16    PCS (imputed)
4    PCS12_adj    Num       8     24    PCS (imputed), adjusted
1    id           Num       8      0
```

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                 19
***  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***             17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

*   X._testimp contains Veterans SF-12 PCS & MCS scores (with MRE imputation)   *
```

Obs	id	IMPUTE_P	PCS12	PCS12_ adj	IMPUTE_M	MCS12	MCS12_ adj
1	32	1	39.8724	40.0294	1	50.9526	51.0961
2	115	0	30.9716	30.8030	0	29.7877	29.3855
3	333	0	30.8019	30.6276	0	40.6368	40.5143
4	379	0	47.5808	47.9670	0	51.3589	51.5127
5	587	0	48.4751	48.8912	0	47.6798	47.7388
6	596	0	40.9062	41.0694	0	20.0427	19.3893
7	861	0	34.8766	34.8384	0	44.3913	44.3655
8	867	0	24.4103	24.0225	0	26.8438	26.3657
9	942	1	37.3533	37.3980	1	36.4930	36.1470
10	1030	0	39.9504	40.0818	0	29.6144	29.2078
11	1098	1	25.8756	24.7227	1	18.7043	16.3184
12	1340	0	45.5016	45.8184	0	53.8011	54.0179
13	1400	0	22.4789	22.0265	0	25.2126	24.6925
14	1789	0	24.2425	23.8490	0	33.2085	32.8945
15	1842	0	24.4999	24.1150	0	24.0159	23.4649
16	2041	0	55.0589	55.6950	0	60.2627	60.6461
17	2118	0	40.5784	40.7307	0	49.6053	49.7140
18	2123	1	24.1250	22.7732	1	47.5376	47.7296
19	2274	0	43.9542	44.2193	0	57.6118	57.9269
20	2353	0	46.8693	47.2317	0	60.0088	60.3857
21	2722	0	30.3115	30.1208	0	23.0119	22.4350
22	2793	0	52.5377	53.0895	0	35.1691	34.9057
23	3075	1	52.8907	54.1696	1	49.3074	49.6199
24	3296	1	36.4142	36.4590	1	54.9967	55.8556
25	3387	0	39.7353	39.8594	0	46.2290	46.2506

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                    20
**  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***              17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

*   X._testimp contains Veterans SF-12 PCS & MCS scores (with MRE imputation)   *

The MEANS Procedure
```

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
id		8630	432919.54	249467.26	32.0000000	863467.00
IMPUTE_P	PCS imputed? (1=yes)	8629	0.2223896	0.4158756	0	1.0000000
PCS12	PCS (imputed)	8629	35.6140090	11.8165637	6.5590801	67.2942192
PCS12_adj	PCS (imputed), adjusted	8629	35.5929108	12.2426073	5.5748626	68.3390146
IMPUTE_M	MCS imputed? (1=yes)	8607	0.2204020	0.4145418	0	1.0000000
MCS12	MCS (imputed)	8607	44.9666271	13.3134214	9.6976570	73.7538304
MCS12_adj	MCS (imputed), adjusted	8607	44.9455010	13.6971092	8.7776281	74.4849548

```
The CORR Procedure

   2  Variables:    PCS12_adj MCS12_adj

                                              Simple Statistics
```

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
PCS12_adj	8629	35.59291	12.24261	307131	5.57486	68.33901	PCS (imputed), adjusted
MCS12_adj	8607	44.94550	13.69711	386846	8.77763	74.48495	MCS (imputed), adjusted

```
          Pearson Correlation Coefficients
             Prob > |r| under H0: Rho=0
               Number of Observations
```

	PCS12_adj	MCS12_adj
PCS12_adj PCS (imputed), adjusted	1.00000 8629	0.30024 <.0001 8606
MCS12_adj MCS (imputed), adjusted	0.30024 <.0001 8606	1.00000 8607

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                        22
***  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***              17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

*   X._testimp contains Veterans SF-12 PCS & MCS scores (with MRE imputation)   *

The UNIVARIATE Procedure
Variable:  PCS12_adj  (PCS (imputed), adjusted)

                            Moments

N                         8629    Sum Weights               8629
Mean                35.5929108    Sum Observations    307131.227
Std Deviation       12.2426073    Variance            149.881432
Skewness             0.2081671    Kurtosis            -0.9302944
Uncorrected SS      12224871.4    Corrected SS           1293177
Coeff Variation     34.3961957    Std Error Mean      0.13179341


              Basic Statistical Measures

    Location                    Variability

Mean     35.59291     Std Deviation           12.24261
Median   34.36438     Variance               149.88143
Mode     55.69496     Range                   62.76415
                      Interquartile Range     19.62849

Quantile       Estimate

100% Max       68.33901
99%            59.63020
95%            56.05356
90%            53.85779
75% Q3         45.17490
50% Median     34.36438
25% Q1         25.54642
10%            20.49750
5%             18.08091
1%             12.78252
0% Min          5.57486
```

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                          23
***  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***                  17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

*   X._testimp contains Veterans SF-12 PCS & MCS scores (with MRE imputation)   *

The UNIVARIATE Procedure
Variable:  MCS12_adj  (MCS (imputed), adjusted)

                            Moments

N                        8607    Sum Weights              8607
Mean                44.945501    Sum Observations   386845.927
Std Deviation      13.6971092    Variance           187.610801
Skewness           -0.2703581    Kurtosis           -0.9230928
Uncorrected SS     19001562.5    Corrected SS       1614578.55
Coeff Variation    30.4749283    Std Error Mean     0.14763965


              Basic Statistical Measures

    Location                    Variability

Mean     44.94550     Std Deviation           13.69711
Median   45.75337     Variance               187.61080
Mode     60.64609     Range                   65.70733
                      Interquartile Range     22.58860

Quantile       Estimate

100% Max       74.48495
99%            67.78900
95%            63.92904
90%            62.10613
75% Q3         57.10104
50% Median     45.75337
25% Q1         34.51244
10%            25.29775
5%             20.91010
1%             16.07151
0% Min          8.77763
```

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                         25
***  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***             17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

Number imputed for PCS and MCS

The FREQ Procedure

Table of IMPUTE_P by IMPUTE_M

IMPUTE_P(PCS imputed? (1=yes))
          IMPUTE_M(MCS imputed? (1=yes))
Frequency,
Percent  ,
Row Pct  ,
Col Pct  ,       .,       0,       1,  Total
ƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆ
       . ,      0 ,      0 ,      1 ,      1
         ,   0.00 ,   0.00 ,   0.01 ,   0.01
         ,   0.00 ,   0.00 , 100.00 ,
         ,   0.00 ,   0.00 ,   0.05 ,
ƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆ
       0 ,      0 ,   6710 ,      0 ,   6710
         ,   0.00 ,  77.75 ,   0.00 ,  77.75
         ,   0.00 , 100.00 ,   0.00 ,
         ,   0.00 , 100.00 ,   0.00 ,
ƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆ
       1 ,     23 ,      0 ,   1896 ,   1919
         ,   0.27 ,   0.00 ,  21.97 ,  22.24
         ,   1.20 ,   0.00 ,  98.80 ,
         , 100.00 ,   0.00 ,  99.95 ,
ƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒˆ
Total          23     6710     1897     8630
             0.27    77.75    21.98   100.00
```

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                       27
***  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***             17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

*  Validity 1: Correlation between PCS & MCS should be low ...  *

The CORR Procedure

   2  Variables:    PCS12_adj MCS12_adj


                                                  Simple Statistics

Variable            N          Mean       Std Dev           Sum       Minimum       Maximum    Label

PCS12_adj        8629      35.59291      12.24261        307131       5.57486      68.33901    PCS (imputed),
adjusted
MCS12_adj        8607      44.94550      13.69711        386846       8.77763      74.48495    MCS (imputed),
adjusted


          Pearson Correlation Coefficients
             Prob > |r| under H0: Rho=0
               Number of Observations

                             PCS12_adj      MCS12_adj

PCS12_adj                      1.00000        0.30024
PCS (imputed), adjusted                        <.0001
                                  8629           8606

MCS12_adj                      0.30024        1.00000
MCS (imputed), adjusted         <.0001
                                  8606           8607
```

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                    28
***  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***              17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

*  Validity 2: PF, RP, and BP items should correlate highest with PCS
 & SF, RE, and MH should correlate highest with MCS  *

The CORR Procedure


              Pearson Correlation Coefficients
                 Prob > |r| under H0: Rho=0
                   Number of Observations

                                     PCS12_adj      MCS12_adj

PF02                                   0.80963        0.33746
Q2b: Moderate Activities                <.0001         <.0001
                                          8515           8496

PF04                                   0.79481        0.31427
Q2d: Climb >1 flights of stairs         <.0001         <.0001
                                          8479           8462

VRP2                                  -0.80623       -0.51417
Q3b: Accomplished less (phys)           <.0001         <.0001
                                          8448           8437

VRP3                                  -0.83188       -0.47852
Q3c: Kind of activities (phys)          <.0001         <.0001
                                          8416           8405

BP2                                   -0.78906       -0.53494
Q7:  Pain interfered with work          <.0001         <.0001
                                          8447           8440

GH1                                   -0.73513       -0.51836
Q1:  Health In General                  <.0001         <.0001
                                          7866           7848
```

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                       29
***  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***                 17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

*  Validity 2: PF, RP, and BP items should correlate highest with PCS
 & SF, RE, and MH should correlate highest with MCS  *

The CORR Procedure

              Pearson Correlation Coefficients
                 Prob > |r| under H0: Rho=0
                   Number of Observations

                                       PCS12_adj      MCS12_adj

VT2                                     -0.62442       -0.61712
Q8e: Lots of energy                       <.0001         <.0001
                                            8467           8467

SF2                                      0.56403        0.76774
Q9:  Time health interfered               <.0001         <.0001
                                            8088           8083

VRE2                                    -0.45234       -0.82051
Q4b: Accomplished less (emot)             <.0001         <.0001
                                            8407           8407

VRE3                                    -0.42712       -0.79534
Q4c: Not as careful as usual              <.0001         <.0001
                                            8378           8378

MH3                                     -0.31447       -0.77654
Q8d: Calm amd peaceful                    <.0001         <.0001
                                            8461           8461

MH4                                      0.21045        0.80290
Q8f: Downhearted and blue                 <.0001         <.0001
                                            8451           8451
```

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                          30
***  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***               17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

*  Validity 3a: Correlations among PCS & MCS scores WITHOUT imputation

The CORR Procedure

   2  Variables:    PCS12_adj MCS12_adj


                                                        Simple Statistics

Variable        N         Mean       Std Dev           Sum       Minimum       Maximum    Label

PCS12_adj    6710     36.04544      12.43784        241865       5.57486      68.33901    PCS (imputed), adjusted
MCS12_adj    6710     45.41924      13.75748        304763       8.77763      74.48495    MCS (imputed), adjusted


     Pearson Correlation Coefficients, N = 6710
             Prob > |r| under H0: Rho=0

                             PCS12_adj      MCS12_adj

PCS12_adj                      1.00000        0.30832
PCS (imputed), adjusted                        <.0001

MCS12_adj                      0.30832        1.00000
MCS (imputed), adjusted         <.0001
```

```
---  Veterans SF-12 imputation program:  v-sf12-impute1.2 , Version 1.2, Sept 2004  ---                          31
***  imputing SF-12 PCS & MCS for dataset: test  (min r2 = .6)  ***               17:34 Monday, September 20, 2004
***  using PCS weights from WT.PCS and MCS weights from WT.MCS  ***

*  Validity 3b: Correlations among PCS & MCS scores WITH imputation

The CORR Procedure

   2  Variables:    PCS12_adj MCS12_adj


                                                  Simple Statistics

Variable         N         Mean       Std Dev           Sum       Minimum       Maximum    Label

PCS12_adj     1919     34.01058      11.39665         65266       6.13031      63.45686    PCS (imputed), adjusted
MCS12_adj     1897     43.26982      13.35071         82083      10.98863      74.24174    MCS (imputed), adjusted


          Pearson Correlation Coefficients
             Prob > |r| under H0: Rho=0
               Number of Observations

                             PCS12_adj      MCS12_adj

PCS12_adj                      1.00000        0.25317
PCS (imputed), adjusted                        <.0001
                                  1919           1896

MCS12_adj                      0.25317        1.00000
MCS (imputed), adjusted         <.0001
                                  1896           1897
```

Merge of Original data and imputed Veterans SF-12 17:34 Monday, September 20, 2004 33

Differences among estimated scores

The MEANS Procedure

Variable	Label	N	Mean	Std Error	t Value	Pr > |t|
d_pcs2	PCS(SF36) - PCS12	8266	0.0728734	0.0352860	2.07	0.0389
d_pcs3	PCS(SF36) - PCS12_adj	8266	0.0836194	0.0356843	2.34	0.0191
d_mcs2	MCS(SF36) - MCS12	8266	-0.0453063	0.0360137	-1.26	0.2084
d_mcs3	MCS(SF36) - MCS12_adj	8266	-0.0340538	0.0362920	-0.94	0.3481

Merge of Original data and imputed Veterans SF-12 17:34 Monday, September 20, 2004 34
Correlations among all scores

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
c3pcs	8266	35.83215	12.20227	296189	3.49836	67.91449	STD PHYSICAL COMPONENT SCALE
PCS12	8629	35.61401	11.81656	307313	6.55908	67.29422	PCS (imputed)
PCS12_adj	8629	35.59291	12.24261	307131	5.57486	68.33901	PCS (imputed), adjusted
c3mcs	8266	45.05498	13.71923	372424	4.82948	76.39240	STD MENTAL COMPONENT SCALE
MCS12	8607	44.96663	13.31342	387028	9.69766	73.75383	MCS (imputed)
MCS12_adj	8607	44.94550	13.69711	386846	8.77763	74.48495	MCS (imputed), adjusted

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	c3pcs	PCS12	PCS12_adj	c3mcs	MCS12	MCS12_adj
c3pcs	1.00000	0.96484	0.96480	0.24799	0.28607	0.28600
STD PHYSICAL COMPONENT SCALE		<.0001	<.0001	<.0001	<.0001	<.0001
	8266	8266	8266	8266	8266	8266
PCS12	0.96484	1.00000	0.99993	0.29042	0.30016	0.30024
PCS (imputed)	<.0001		<.0001	<.0001	<.0001	<.0001
	8266	8629	8629	8266	8606	8606
PCS12_adj	0.96480	0.99993	1.00000	0.29035	0.30014	0.30024
PCS (imputed), adjusted	<.0001	<.0001		<.0001	<.0001	<.0001
	8266	8629	8629	8266	8606	8606
c3mcs	0.24799	0.29042	0.29035	1.00000	0.97110	0.97103
STD MENTAL COMPONENT SCALE	<.0001	<.0001	<.0001		<.0001	<.0001
	8266	8266	8266	8266	8266	8266
MCS12	0.28607	0.30016	0.30014	0.97110	1.00000	0.99991
MCS (imputed)	<.0001	<.0001	<.0001	<.0001		<.0001
	8266	8606	8606	8266	8607	8607
MCS12_adj	0.28600	0.30024	0.30024	0.97103	0.99991	1.00000
MCS (imputed), adjusted	<.0001	<.0001	<.0001	<.0001	<.0001	
	8266	8606	8606	8266	8607	8607

Appendix C: The SAS© Macro for Imputation of Veterans SF-12

```
 ********************************************************************
 *  PROGRAM: v-sf12-impute1.2.sas     September 2004                *
 *                                                                  *
 *  Veterans SF-12 imputation program, using weights based on the VA  *
 *    1999 Large Health Survey of Veteran Enrollees                  *
 *                                                                  *
 *  Note: Program computes adjusted and unadjusted PCS and MCS scores *
 *    but reports only adjusted scores (corrected for regression to   *
 *    the mean).                                                     *
 ********************************************************************;

%macro VSF12IMP (indata=, idvar= id, minr2= .6, OMIT=1,
                 PCS_WTS=, MCS_WTS=,
                 validity=0, outdata= _imputed);

  %let PROG = %str(v-sf12-impute1.2);
  %let VER  = %str(Version 1.2, Sept 2004);

 /***************************************************************************
 *  To INCLUDE the macro in a SAS program use the following statement:
 *
 *    %include '<source>\v-sf12-impute.sas';  where <source> is a path, e.g.,
 *         c:\sf12
 *
 *  To EXECUTE the macro, once the code is included in your SAS program,
 *   include the following statement:
 *
 *    %macro VSF12IMP(indata= , idvar= id, minr2= .6, PCS_WTS=, MCS_WTS=,
 *       validity=0, outdata= _imputed );
 *
 * where you specify values for (or use the defaults):
 *
 * indata     name of the input sas dataset (REQUIRED)
 * idvar      name of the SAS variable idenfying each case uniquely [default= ID]
 * minr2      a number from 0 to 1, specifying the minimum value of R2 to allow
 *            in the models predicting PCS & MCS from subsets of SF12 responses
 *            [default = .6]
 * omit       Remove cases with ALL SF-12 items missing? 0=NO 1=Yes [default]
 * pcs_wts    names of SAS datasets containing the 4096 sets of weights to predict
 * mcs_wts    PCS & MCS from subsets of SF12 responses [REQUIRED]
 * validity   1 = do checks on scoring validity, 0 [default] = ignore checks
 * outdata    name of output dataset containing imputed PCS/MCS scores [default=
_imputed]
 *
 *
 * NOTES:
 *   1) The dataset named by INDATA must include a case identification
 *          variable (IDVAR) & be sorted by that variable
 *   2) The SF12 items must be numeric (not character) variables and be named:
 *             pf02 pf04 vrp2 vrp3 bp2 gh1 vt2 sf2 vre2 vre3 mh3 mh4
 ***************************************************************************/
```

```
title "---  Veterans SF-12 imputation program:  &PROG , &VER  ---";
title2 "***  imputing SF-12 PCS & MCS for dataset: &INDATA  (min r2 = &minr2)
***";
title3 "***  using PCS weights from &PCS_WTS and MCS weights from &MCS_WTS  ***";

data _null_;
 put /// '******************************************************************';
 put  @10 ' Veterans SF-12 Imputation Program for HOS' /
            @12 ' Health Outcomes Technologies Program' /
            @12 ' Boston University School of Public Health' /
            @12 ' Program Version 1.1, September 2004' //
       '         Supported by NCQA/CMS, Boston University, and' /
       ' the Research Services of the US Department of Veterans Affairs';

 put / @15 "Name of dataset for analysis:  &indata" /
       @15 "Case identifier:               &idvar" /
       @15 "Minimum R2 for imputation:     &minr2" //
       @15 "PCS weights are read from:     &PCS_WTS" /
       @15 "MCS weights are read from:     &MCS_WTS" /;
 if &validity=1 then put  @15 'Validity check is: ON' /;
  else put  @15 'Validity check is: OFF' /;
 if &omit=1 then put  @15 'Cases with all SF-12 items missing are:  DELETED' / ;
  else put  @15 'Cases with all SF-12 items missing are:  INCLUDED' / ;

 put '******************************************************************';


*********************************************************************
*  Error checks. If fail, then DO NOT run the macro (print warnings)*
*********************************************************************;

 /* check that required parameters are present and valid */
 %if %length(&PCS_WTS)=0 |  %length(&MCS_WTS)=0 %then %do;
  put // " ---> Error:  Filenames not specified for PCS_WTS and MCS_WTS " //
         " ***> PROGRAM WILL TERMINATE. Next time, specify names for PCS_WTS &
MCS_WTS parameters" ////;

 %end;
 %else %if %length(&indata)=0 %then %do;
  put // " ---> Error:  Filename not specified for input dataset containing SF12
items " //
         " ***> PROGRAM WILL TERMINATE. Next time, specify names for INDATA
parameter  " ////;
 %end;
 %else %if (.<&minr2<0 | &minr2>1) | %length (&minr2)=0 %then %do;
  put // " ---> Error:  Value for minimum r2 parameter (&minr2) is out of bounds
(0, 1) " //
         " ***> PROGRAM WILL TERMINATE. " ////;
 %end;
 %else %do; /* If parameters are OK, then run the MAJOR loop */

*******************************************************************
***                     0. input data                           ***
*******************************************************************;

*********************************************************************
*  Read in the INDATA file, and keep only ID variable & SF12 items  *
```

```
*******************************************************************;

data _sf12scal;
 length pf02 pf04 vrp2 vrp3 bp2 gh1 vt2 sf2 vre2 vre3 mh3 mh4 3;
 set &INDATA (keep=&IDVAR pf02 pf04 vrp2 vrp3 bp2 gh1 vt2 sf2 vre2 vre3 mh3 mh4);

 * optional removal of cases with all items missing;
  if &omit=1 then do;
    if n(of pf02 pf04 vrp2 vrp3 bp2 gh1 vt2 sf2 vre2 vre3 mh3 mh4)=0 then DELETE;
  end;

******************************************************************
***              step 1: data cleaning                        ***
***  change out-of-range values to missing for each item.       ***
*******************************************************************;

array threept pf02 pf04;
  do over threept;
  if threept lt 1 or threept gt 3 then threept = .;
end;

array fivept  vrp2 vrp3 vre2 vre3 bp2 sf2 gh1;
  do over fivept;
  if fivept lt 1 or fivept gt 5 then fivept = .;
end;

array sixpt mh3 mh4 vt2;
  do over sixpt;
  if sixpt lt 1 or sixpt gt 6 then sixpt = .;
end;

******************************************************************
*                step 2: create 47 indicator variables from      *
*                        item response choices                   *
*******************************************************************;
length  pf02_2 pf02_3 pf04_2 pf04_3 vrp2_2 vrp2_3 vrp2_4
      vrp2_5 vrp3_2 vrp3_3 vrp3_4 vrp3_5 vre2_2 vre2_3
      vre2_4 vre2_5 vre3_2 vre3_3 vre3_4 vre3_5 bp2_2
      bp2_3  bp2_4  bp2_5  mh3_2  mh3_3  mh3_4  mh3_5
      mh3_6  mh4_2  mh4_3  mh4_4  mh4_5  mh4_6  vt2_2
      vt2_3  vt2_4  vt2_5  vt2_6  sf2_2  sf2_3  sf2_4
      sf2_5  gh1_2  gh1_3  gh1_4  gh1_5  3;

pf02_2 = .;
  if pf02 = . then pf02_2 = .; else
  if pf02 = 2 then pf02_2 = 1; else pf02_2 = 0;

pf02_3 = .;
  if pf02 = . then pf02_3 = .; else
  if pf02 = 3 then pf02_3 = 1; else pf02_3 = 0;

pf04_2 = .;
  if pf04 = . then pf04_2 = .; else
  if pf04 = 2 then pf04_2 = 1; else pf04_2 = 0;

pf04_3 = .;
  if pf04 = . then pf04_3 = .; else
```

```
  if pf04 = 3 then pf04_3 = 1; else pf04_3 = 0;

vrp2_2 = .;
  if vrp2 = . then vrp2_2 = .; else
  if vrp2 = 2 then vrp2_2 = 1; else vrp2_2 = 0;

vrp2_3 = .;
  if vrp2 = . then vrp2_3 = .; else
  if vrp2 = 3 then vrp2_3 = 1; else vrp2_3 = 0;

vrp2_4 = .;
  if vrp2 = . then vrp2_4 = .; else
  if vrp2 = 4 then vrp2_4 = 1; else vrp2_4 = 0;

vrp2_5 = .;
  if vrp2 = . then vrp2_5 = .; else
  if vrp2 = 5 then vrp2_5 = 1; else vrp2_5 = 0;

vrp3_2 = .;
  if vrp3 = . then vrp3_2 = .; else
  if vrp3 = 2 then vrp3_2 = 1; else vrp3_2 = 0;

vrp3_3 = .;
  if vrp3 = . then vrp3_3 = .; else
  if vrp3 = 3 then vrp3_3 = 1; else vrp3_3 = 0;

vrp3_4 = .;
  if vrp3 = . then vrp3_4 = .; else
  if vrp3 = 4 then vrp3_4 = 1; else vrp3_4 = 0;

vrp3_5 = .;
  if vrp3 = . then vrp3_5 = .; else
  if vrp3 = 5 then vrp3_5 = 1; else vrp3_5 = 0;

bp2_2 = .;
  if bp2 = . then bp2_2 = .; else
  if bp2 = 2 then bp2_2 = 1; else bp2_2 = 0;

bp2_3 = .;
  if bp2 = . then bp2_3 = .; else
  if bp2 = 3 then bp2_3 = 1; else bp2_3 = 0;

bp2_4 = .;
  if bp2 = . then bp2_4 = .; else
  if bp2 = 4 then bp2_4 = 1; else bp2_4 = 0;

bp2_5 = .;
  if bp2 = . then bp2_5 = .; else
  if bp2 = 5 then bp2_5 = 1; else bp2_5 = 0;

gh1_2 = .;
  if gh1 = . then gh1_2 = .; else
  if gh1 = 2 then gh1_2 = 1; else gh1_2 = 0;

gh1_3 = .;
  if gh1 = . then gh1_3 = .; else
  if gh1 = 3 then gh1_3 = 1; else gh1_3 = 0;
```

```
gh1_4 = .;
  if gh1 = . then gh1_4 = .; else
  if gh1 = 4 then gh1_4 = 1; else gh1_4 = 0;

gh1_5 = .;
  if gh1 = . then gh1_5 = .; else
  if gh1 = 5 then gh1_5 = 1; else gh1_5 = 0;

vt2_2 = .;
  if vt2 = . then vt2_2 = .; else
  if vt2 = 2 then vt2_2 = 1; else vt2_2 = 0;

vt2_3 = .;
  if vt2 = . then vt2_3 = .; else
  if vt2 = 3 then vt2_3 = 1; else vt2_3 = 0;

vt2_4 = .;
  if vt2 = . then vt2_4 = .; else
  if vt2 = 4 then vt2_4 = 1; else vt2_4 = 0;

vt2_5 = .;
  if vt2 = . then vt2_5 = .; else
  if vt2 = 5 then vt2_5 = 1; else vt2_5 = 0;

vt2_6 = .;
  if vt2 = . then vt2_6 = .; else
  if vt2 = 6 then vt2_6 = 1; else vt2_6 = 0;

sf2_2 = .;
  if sf2 = . then sf2_2 = .; else
  if sf2 = 2 then sf2_2 = 1; else sf2_2 = 0;

sf2_3 = .;
  if sf2 = . then sf2_3 = .; else
  if sf2 = 3 then sf2_3 = 1; else sf2_3 = 0;

sf2_4 = .;
  if sf2 = . then sf2_4 = .; else
  if sf2 = 4 then sf2_4 = 1; else sf2_4 = 0;

sf2_5 = .;
  if sf2 = . then sf2_5 = .; else
  if sf2 = 5 then sf2_5 = 1; else sf2_5 = 0;

vre2_2 = .;
  if vre2 = . then vre2_2 = .; else
  if vre2 = 2 then vre2_2 = 1; else vre2_2 = 0;

vre2_3 = .;
  if vre2 = . then vre2_3 = .; else
  if vre2 = 3 then vre2_3 = 1; else vre2_3 = 0;

vre2_4 = .;
  if vre2 = . then vre2_4 = .; else
  if vre2 = 4 then vre2_4 = 1; else vre2_4 = 0;
```

```
vre2_5 = .;
  if vre2 = . then vre2_5 = .; else
  if vre2 = 5 then vre2_5 = 1; else vre2_5 = 0;

vre3_2 = .;
  if vre3 = . then vre3_2 = .; else
  if vre3 = 2 then vre3_2 = 1; else vre3_2 = 0;

vre3_3 = .;
  if vre3 = . then vre3_3 = .; else
  if vre3 = 3 then vre3_3 = 1; else vre3_3 = 0;

vre3_4 = .;
  if vre3 = . then vre3_4 = .; else
  if vre3 = 4 then vre3_4 = 1; else vre3_4 = 0;

vre3_5 = .;
  if vre3 = . then vre3_5 = .; else
  if vre3 = 5 then vre3_5 = 1; else vre3_5 = 0;

mh3_2 = .;
  if mh3 = . then mh3_2 = .; else
  if mh3 = 2 then mh3_2 = 1; else mh3_2 = 0;

mh3_3 = .;
  if mh3 = . then mh3_3 = .; else
  if mh3 = 3 then mh3_3 = 1; else mh3_3 = 0;

mh3_4 = .;
  if mh3 = . then mh3_4 = .; else
  if mh3 = 4 then mh3_4 = 1; else mh3_4 = 0;

mh3_5 = .;
  if mh3 = . then mh3_5 = .; else
  if mh3 = 5 then mh3_5 = 1; else mh3_5 = 0;

mh3_6 = .;
  if mh3 = . then mh3_6 = .; else
  if mh3 = 6 then mh3_6 = 1; else mh3_6 = 0;

mh4_2 = .;
  if mh4 = . then mh4_2 = .; else
  if mh4 = 2 then mh4_2 = 1; else mh4_2 = 0;

mh4_3 = .;
  if mh4 = . then mh4_3 = .; else
  if mh4 = 3 then mh4_3 = 1; else mh4_3 = 0;

mh4_4 = .;
  if mh4 = . then mh4_4 = .; else
  if mh4 = 4 then mh4_4 = 1; else mh4_4 = 0;

mh4_5 = .;
  if mh4 = . then mh4_5 = .; else
  if mh4 = 5 then mh4_5 = 1; else mh4_5 = 0;

mh4_6 = .;
```

```
  if mh4 = . then mh4_6 = .; else
  if mh4 = 6 then mh4_6 = 1; else mh4_6 = 0;


*******************************************************************
*                step 3: Create the "number" variable for         *
*                        the observed data, based on the pattern  *
*                        of observed indicators. Then sort        *
*                                                                 *
* NE1 contains: &idvar, number, 47 SF12 response indicators       *
*******************************************************************;
data _ne1 (DROP = pf02 pf04 vrp2 vrp3 bp2 gh1 vt2 sf2 vre2 vre3 mh3 mh4 I A);
 set _sf12scal;

* number is an index variable ranging from 0 to 4095 which indicates the
  observed pattern of missing Sf12 items in the data. 0 is number for
  complete SF12 items, 1 has last item (MH4) missing, 2 has next last item (MH3)
missing,
  3 has last 2 items both missing, 4 has 3rd last item (Re3) missing, 5 has 3rd
and last missing,
  6 has last 2nd and 3rd last missing, 7 has last 3 all missing, etc., up to 4095
which has
  all items missing ;
number=0;
array sf36i(12) pf02 pf04 vrp2 vrp3 bp2 gh1 vt2 sf2 vre2 vre3 mh3 mh4;
do i=1 to 12;
  number = number*2;
  if sf36i(i)=. then number = number +1;
end;

array sf12v(47) pf02_2 pf02_3 pf04_2 pf04_3 vrp2_2 vrp2_3 vrp2_4
            vrp2_5 vrp3_2 vrp3_3 vrp3_4 vrp3_5 vre2_2 vre2_3
            vre2_4 vre2_5 vre3_2 vre3_3 vre3_4 vre3_5 bp2_2
            bp2_3  bp2_4  bp2_5  mh3_2  mh3_3  mh3_4  mh3_5
            mh3_6  mh4_2  mh4_3  mh4_4  mh4_5  mh4_6  vt2_2
            vt2_3  vt2_4  vt2_5  vt2_6  sf2_2  sf2_3  sf2_4
            sf2_5  gh1_2  gh1_3  gh1_4  gh1_5;
do a=1 to 47;
  if sf12v(a)=. then sf12v(a)=0;
end;

proc sort data=_NE1;
 by number;

*********************************************************************
*                step 4: weighting and aggregation of               *
*                        indicator variables using                  *
*                        physical and mental regression weights, with*
*                  missing value imputation included       *
*********************************************************************
*  4a. Impute PCS scores
*********************************************************************;
* Select certain PCS imputation models, based on r2 value greater than MINR2;
data _pcsuse;
 set &PCS_WTS;
  if r2>= &minr2;
```

```
 label number='Index for imputation model (0-4095)'
  r2='R2 of regression model for index model'
  items='# of valid items for index model' ;

* Impute PCS scores;
data _PCSI (drop = Bpf2r2--Bmh4r6 pf02_2--gh1_5 cons);
 merge _ne1(in=PP) _pcsuse(in=PU);
  by number;
   if PP & PU;  /* Keep patterns IFF in BOTH NE1 & in PCSUSE */

 IF NUMBER = 0 THEN IMPUTE_P = 0; ELSE IMPUTE_P = 1;
 LABEL IMPUTE_P = 'PCS imputed? (1=yes)';

  PCS12 =pf02_2 *Bpf2r2 + pf02_3*Bpf2r3  + pf04_2*Bpf4r2  + pf04_3*Bpf4r3
     +vrp2_2*Bvrp2r2+ vrp2_3*Bvrp2r3 + vrp2_4*Bvrp2r4 + vrp2_5*Bvrp2r5
     +vrp3_2*Bvrp3r2+ vrp3_3*Bvrp3r3 + vrp3_4*Bvrp3r4 + vrp3_5*Bvrp3r5
     +vre2_2*Bvre2r2+ vre2_3*Bvre2r3 + vre2_4*Bvre2r4 + vre2_5*Bvre2r5
     +vre3_2*Bvre3r2+ vre3_3*Bvre3r3 + vre3_4*Bvre3r4 + vre3_5*Bvre3r5
     +bp2_2 *Bbp2r2 + bp2_3 *Bbp2r3  + bp2_4 *Bbp2r4  + bp2_5 *Bbp2r5
     +mh3_2 *Bmh3r2 + mh3_3 *Bmh3r3  + mh3_4 *Bmh3r4  + mh3_5 *Bmh3r5
     +mh3_6 *Bmh3r6 + mh4_2 *Bmh4r2  + mh4_3 *Bmh4r3  + mh4_4 *Bmh4r4
     +mh4_5 *Bmh4r5 + mh4_6 *Bmh4r6  + vt2_2 *Bvt2r2  + vt2_3 *Bvt2r3
     +vt2_4 *Bvt2r4 + vt2_5 *Bvt2r5  + vt2_6 *Bvt2r6  + sf2_2 *Bsf2r2
     +sf2_3 *Bsf2r3 + sf2_4 *Bsf2r4  + sf2_5 *Bsf2r5  + gh1_2 *Bgh1r2
     +gh1_3 *Bgh1r3 + gh1_4 *Bgh1r4  + gh1_5 *Bgh1r5  + cons;

 IF R2>0 THEN PCS12_adj=36.02+(PCS12-36.02)/(sqrt(r2));

  label PCS12='PCS (imputed)' PCS12_adj='PCS (imputed), adjusted';

*******************************************************************;
*  4b. Impute MCS scores
*******************************************************************;
* Select certain MCS imputation models, based on r2 value greater than MINR2;
data _mcsuse;
 set &MCS_WTS;
  if r2>=&minr2;

 label number='Index for imputation model (0-4095)'
  r2='R2 of regression model for index model'
  items='# of valid items for index model' ;

* Impute MCS scores;
data _MCSI (drop = Bpf2r2--Bmh4r6 pf02_2--gh1_5 cons);
 merge _ne1 (in=PP) _mcsuse (in=MU);
  by number;
   if PP & MU;  /* Keep patterns IFF in BOTH NE1 & in MCSUSE */

 IF NUMBER = 0 THEN IMPUTE_M = 0; ELSE IMPUTE_M = 1;
 LABEL IMPUTE_M = 'MCS imputed? (1=yes)';

 MCS12 =  pf02_2 *Bpf2r2 + pf02_3*Bpf2r3  + pf04_2*Bpf4r2  + pf04_3*Bpf4r3
     +vrp2_2*Bvrp2r2+ vrp2_3*Bvrp2r3 + vrp2_4*Bvrp2r4 + vrp2_5*Bvrp2r5
     +vrp3_2*Bvrp3r2+ vrp3_3*Bvrp3r3 + vrp3_4*Bvrp3r4 + vrp3_5*Bvrp3r5
     +vre2_2*Bvre2r2+ vre2_3*Bvre2r3 + vre2_4*Bvre2r4 + vre2_5*Bvre2r5
     +vre3_2*Bvre3r2+ vre3_3*Bvre3r3 + vre3_4*Bvre3r4 + vre3_5*Bvre3r5
     +bp2_2 *Bbp2r2 + bp2_3 *Bbp2r3  + bp2_4 *Bbp2r4  + bp2_5 *Bbp2r5
```

```
     +mh3_2 *Bmh3r2 + mh3_3 *Bmh3r3  + mh3_4 *Bmh3r4  + mh3_5 *Bmh3r5
     +mh3_6 *Bmh3r6 + mh4_2 *Bmh4r2  + mh4_3 *Bmh4r3  + mh4_4 *Bmh4r4
     +mh4_5 *Bmh4r5 + mh4_6 *Bmh4r6  + vt2_2 *Bvt2r2  + vt2_3 *Bvt2r3
     +vt2_4 *Bvt2r4 + vt2_5 *Bvt2r5  + vt2_6 *Bvt2r6  + sf2_2 *Bsf2r2
     +sf2_3 *Bsf2r3 + sf2_4 *Bsf2r4  + sf2_5 *Bsf2r5  + gh1_2 *Bgh1r2
     +gh1_3 *Bgh1r3 + gh1_4 *Bgh1r4  + gh1_5 *Bgh1r5  + cons;

 IF R2>0 THEN   MCS12_adj=45.39+(MCS12-45.39)/(sqrt(r2));

  label MCS12='MCS (imputed)' MCS12_adj='MCS (imputed), adjusted';

********************************************************************
***              step 5: Combine imputed scores into 1 file    ***
********************************************************************;
proc sort data=_PCSI; by &idvar;
proc sort data=_MCSI; by &idvar;

data &outdata(drop=number items r2);
 merge _pcsi _mcsi;
  by &idvar;

*-----------------------------------------------------------------*;
proc contents;
  title5 "*   &outdata contains Veterans SF-12 PCS & MCS scores (with MRE
imputation)   *";

proc print data= &outdata (obs=25);
proc means;

proc corr;
  var PCS12_adj MCS12_adj;
proc univariate ;
  var PCS12_adj MCS12_adj;

proc freq;
  title5 'Number imputed for PCS and MCS';
  tables impute_p*impute_m / missing;

********************************************************************
***              step 6: cleanup data sets                     ***
********************************************************************;

proc datasets NOlist; delete _ne1 _pcsuse _mcsuse _sf12scal _pcsi _mcsi;

********************************************************************
***              step 7: optional validity check               ***
***     based on SF12 scoring manual, p. 13                    ***
********************************************************************;

%if &validity=1 %then %do;
 data _val;
  merge &outdata (keep= &idvar PCS12 PCS12_adj MCS12 MCS12_adj impute_m impute_p)
        &indata;
   by &idvar;

 proc corr;
   title5 '*  Validity 1: Correlation between PCS & MCS should be low ...  *';
```

```
512      var PCS12_adj MCS12_adj;
513
514    proc corr;
515      title5 '*  Validity 2: PF, RP, and BP items should correlate highest with PCS';
516      title6 ' & SF, RE, and MH should correlate highest with MCS  *';
517      var PCS12_adj MCS12_adj;
518      with pf02 pf04 vrp2 vrp3 bp2 gh1 vt2 sf2 vre2 vre3 mh3 mh4 ;
519
520    proc corr;
521     title5 '*  Validity 3a: Correlations among PCS & MCS scores WITHOUT imputation';
522     where IMPUTE_P = 0 & IMPUTE_M = 0;
523     var  PCS12_adj MCS12_adj;
524    proc corr;
525     title5 '*  Validity 3b: Correlations among PCS & MCS scores WITH imputation';
526     where IMPUTE_P = 1 | IMPUTE_M = 1;
527     var  PCS12_adj MCS12_adj;
528
529    proc datasets NOlist; delete _val;
530   %end;  /* end validity loop */
531
532   %end;  /* end MAJOR loop */
533
534    title4;
535    data _null_ ;
536      put ///
537   '*****************************************************************' /
538   '       --> End of Veterans SF-12 Imputation Program <--' /
539   '*****************************************************************' //;
540
541   %MEND;
542   *********************************************************************;
543
```

Comments on the Sample Program

Line numbers	Comments
1-12	description
14-16	Begin the macro and define the required & default variables
18-19	Identify the program version
21-52	Instructions for use
56-80	Titles and header information printed to log file
83-105	Error checks
107-22	Read the input data (keep only case identifier and the SF-12 items), and optionally, omit if all SF-12 items are missing
124-142	Data cleaning (if item responses are out of range, then set them to missing)
144-342	Define and create the 47 indicator variables for the 12 items of the Veterans SF-12 (value 1 is the omitted level for each SF-12 item)
345-83	For the dataset input, compute “number” (0 – 4095) which indicates the pattern of missing data (0 -> all 12 items are present), and then sort the data by “number”

Line numbers	Comments
385-400	Read in the dataset of PCS weights, if the r2 value for a model exceeds the minimum r2 specified by the user
402-426	Merge the dataset of PCS weights with the input dataset, by “number”, if a given pattern of data is both observed in the user’s data and in the set of PCS weights. Then compute the imputed PCS score (PCS12), and the adjusted PCS score (PCS12_ADJ), which is adjusted by the square root of the r2 for that pattern of data
428-438	Repeat for MCS weights
440-464	Repeat to estimate MCS scores
466-474	Combine the imputed PCS and MCS scores in dataset specified by the “outdata” option on the macro statement.
476-491	Print selected results, including a cross-tab identifying whether a score was imputed for PCS or MCS …
493-497	Delete intermediate datasets
499-530	Optionally, combine the input data and the imputed data to conduct some validity checks
532-541	End the macro

Appendix D: SF-12 Questions

1. In general, would you say your health is: **(GH1)**

Excellent	Very good	Good	Fair	Poor
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

2. The following items are about activities you might do during a typical day. Does **your health now limit you** in these activities? If so, how much?

ACTIVITIES	Yes, limited a lot	Yes, limited a little	No, not limited at all
a. **Moderate activities,** such as moving a table, pushing a vacuum cleaner, bowling, or playing golf ....**(PF2)**........	1 ☐	2 ☐	3 ☐
b. Climbing **several** flights of stairs......................**(PF4)**......	1 ☐	2 ☐	3 ☐

3. During the **past 4 weeks,** have you had any of the following problems with your work or other regular daily activities **as a result of your physical health?**

	No, none of the time	Yes, a little of the time	Yes, some of the time	Yes, most of the time	Yes, all of the time
a. **Accomplished less** than you would like...... **(VRP2)**	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
b. Were limited in the **kind** of work or other activities ....**(VRP3)**	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

4. During the **past 4 weeks,** have you had any of the following problems with your work or other regular daily activities **as a result of any emotional problems** (such as feeling depressed or anxious)?

	No, none of the time	Yes, a little of the time	Yes, some of the time	Yes, most of the time	Yes, all of the time
a. **Accomplished less** than you would like...... **(VRE2)**	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
b. Didn't do work or other activities as **carefully** as usual.............................**(VRE3)**	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

5. During the **past 4 weeks,** how much did **pain** interfere with your normal work (including both work outside the home and housework)? **(BP2)**

Not at all	A little bit	Moderately	Quite a bit	Extremely
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

These questions are about how you feel and how things have been with you during the **past 4 weeks.** For each question, please give the one answer that comes closest to the way you have been feeling.

6. How much of the time during the **past 4 weeks:**

	All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	None of the time
a. Have you felt calm and peaceful?..........................**(MH3)**	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
b. Did you have a lot of energy?...... **(VT2)**	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐
c. Have you felt downhearted and blue?..........................**(MH4)**	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐

7. During the **past 4 weeks,** how much of the time has your **physical health or emotional problems** interfered with your social activities (like visiting with friends, relatives, etc.)? **(SF2)**

All of the time	Most of the time	Some of the time	A little of the time	None of the time
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

Now, we'd like to ask you some questions about how your health may have changed.

8. **Compared to one year ago,** how would you rate your **physical health** in general **now?**

Much better	Slightly better	About the same	Slightly worse	Much worse
1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

9. **Compared to one year ago,** how would you rate your **emotional problems** (such as feeling anxious, depressed or irritable) in general **now?**

Much better	Slightly better	About the same	Slightly worse	Much worse
1☐	2☐	3☐	4☐	5☐

APPENDIX E. Evaluation Survey for Alpha Testing of Manual, Users Guide and Computer Program

1. How did you find the manual overall?

   Excellent =1 Very Good =2 Good =3 Fair=4 Poor=5

   a. Did the manual seem to be complete?
      Very complete=1 Complete=2 partially complete=3 and not at all complete=4

   b. Were there any particular sections that were strong?

   c. Were there any particular sections that could be strengthened?

   d. Did the manual include any sections that could be omitted?

   e. Were the theory and methods adequately explained?
      Very well explained=1 to not at all explained=5

   f. Was the scoring methodology for the Veterans SF-12 well articulated?
      Very well articulated=1 to not at all articulated=5

2. Overall, how did you find the users guide for the Veterans SF-12 Imputation program?

   Excellent =1 Very Good =2 Good =3 Fair=4 Poor=5

   a. Was the users guide clearly written?

   Very clearly written=1 to not at all well written=5

   b. How much did the users guide help in the running of the imputation program on the test data set?

   A lot=1 to very little=5

   c. What were the strengths of the users guide for the Veterans SF-12 imputation program?

   d. What were it's weaknesses?

3. In terms of the running of the imputation program on the test data set:

   a. How much time did it take you to read the users guide and run the test data set using the imputation program?

   b. How easy was it to run the program?

   Very easy=1 to very difficult =5

   c. Were there any particular problems you encountered when attempting to run the program?

References

Cooper JK, Kohlmann, T., Michael JA, Haffer SC, Stevic M. Health outcomes. New quality measure for Medicare. International Journal for Quality in Health Care, 2001; 13:9-16.

Embretson SE and Reise SP. *Item Response Theory for Psychologists*. Erlbaum Associates Publishers, London, 2000.

Gandek, B, Sinclair SJ, Kosinski M, & Ware JE (2004). Psychometric evaluation of the SF-36 health survey in Medicare managed care. Health Care Financing Review, 25(4), 5-25.

Haffer SC & Bown SE (2004). measuring and improving health outcomes in Medicare: The Medicare HOS program. Health Care Financing Review, 25(4), 1-3.

Jones N, Jones, SL, Miller NA (2004). The Medicare Health Outcomes Survey program: Overview, context, and near-term prospects. Health & Quality of Life Outcomes, 2, 33. (www.hqlo.com/content/2/1/33)

Kazis LE. The veterans SF-36 health status questionnaire: Development and application in the Veterans Health Administration. Medical Outcomes Trust Monitor 2000; 5 (1).

Kazis LE. et. al. Health Status of Veterans: Physical and Mental Component Summary Scores (SF-12V). 1997 National Survey of Ambulatory Care Patients, Executive Report. Office of Performance and Quality, Health Assessment Project HSR&D Field Program, VHA National Customer Feedback Center, Washington, D.C., Bedford and West Roxbury, Massachusetts, April 1998.

Kazis L, et al. 1999 Large Health Survey of Veteran Enrollees: Health status and outcomes of veterans: Veterans SF-36 physical and mental component summary scores (Executive Report). Office of Quality and Performance, Veterans Health Administration, US Department of Veterans Affairs, Washington, D.C., May 2000.

Kazis LE, Lee, A., Spiro A III, Rogers WH, Ren XS, Miller DR, Selim A, Hamed A, Haffer SC (2004a). Measurement comparisons of the Medical Outcomes Study and Veterans SF-36 Health Survey. Health Care Financing Review, 25(4),43-57.

Kazis, L. E., Miller, D. R., Clark, J., Skinner, K., Lee, A., Ren, X. S., Spiro, A. III, Rogers, W. H., & Ware, J.E. (2004b). Improving the response choices on the Veterans SF-36 health survey role functioning scales: Results from the Veterans Health Study. Journal of Ambulatory Care Management, 27, 263-80.

Kazis LE, Miller DR, Clark J, Skinner K, Lee A, Rogers W, Spiro III. A, Payne S, Fincke G, Selim A, Linzer M. Health related quality of life in patients served by the Department of Veterans Affairs: Results from the Veterans Health Study. Arch Intern Med. 1998;158:626-632.

Kazis LE. et. al. Health Status and Outcomes of Veterans: Physical and Mental Component Summary Scores (Veterans SF-12). 1998 National Survey of Hospitalized Patients, Executive Report. Office of Performance and Quality, Health Assessment Project, HSR&D Field Program, Washington, D.C. and Bedford, Massachusetts, April 1999.

Kazis L., Ren XS, Lee A, et al. Health status in VA patients: Results from the Veterans Health Study. American Journal of Medical Quality 1999; 14(1):28-38.

Kosinski M, Bayliss, M., Bjorner, J.B., & Ware, J.E. (2000). Improving estimates of SF-36 health survey scores for respondents with missing data. Medical Outcomes Trust Monitor Fall 2000; pp 8-10.

Little RJA, Rubin, DB: *Statistical Analysis with Missing Data.* 278 pp. Wiley, New York, 1987.

Rogers W, Qian S and Kazis L. (2003, Nov.). Imputing the Physical and Mental Summary Scores (PCS and MCS) for the MOS SF-36 and the Veterans SF-36 Health Survey in the presence of Missing Data. Report prepared for the National Committee for Quality Assurance under Contract No. 500-00-0055, entitled Implementing the HEDIS Medicare Health Outcomes Survey, sponsored by the Centers for Medicare & Medicaid Services, U.S. Department of Health and Human Services. Final Report July 2004.

Ware JE. Scales for measuring general health perceptions. Health Services Research 1976; 11, 396-415.

Ware JE. Modifications to the Role Functioning Scales of the Veterans SF-36 (Private Communication 1996).

Ware JE, Kosinski M: *SF-36 Physical & Mental Health Summary Scores: A Manual for Users of Version 1. 2nd Edition.* QualityMetric Incorporated, Lincoln RI. 2001.

Ware JE, Kosinski M, Dewey JE. How to Score Version 2 of the SF-36 Health Survey. Medical Outcomes Trust and QualityMetric Incorporated 2002.

Ware JE, Kosinski M, Keller SD. *SF-36 Physical and Mental Health Summary Scales, A User's Manual.* The Health Institute, New England Medical Center. Boston, MA, 1994.

Ware JE, Kosinski M, Keller SD: *How to Score the SF-12 Physical & Mental health Summary Scales.* The Health Institute, New England Medical Center. Boston, MA, 1995.

Ware JE, Kosinski M and Keller SD. A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. Medical Care 1996; 34: 220-233.

Ware JE, Snow KK, Kosinski M, Gandek B: *SF-36 Health Status Survey Manual and Interpretation Guide.* Boston. The Health Institute, New England Medical Center. Boston, MA, 1993.