



SF-36v2[®] Advanced Scoring Guidelines: How to submit data, what you will receive back, and scoring timeline

Optum's goals are to make the process of submitting data as straightforward as possible and make your resulting scored dataset clear, intuitive, and easy to access. Adhering closely to the following instructions will achieve both of these goals.

If you have any questions on this process, please don't hesitate to contact Scoring Help at scoringhelp@qualitymetric.com.

1. These guidelines apply to

- SF-36 (If you have SF-12 data please contact us)
- **v2** data only (If you have SF-36v1[®] data please contact us)
- Both standard (four-week) and acute (one-week) recall periods
- Advanced certified scoring (If you have purchased basic certified scoring, please contact us)

2. Preparing your data to submit to us

- **Structure**

The data must be structured so that each case (each row or line of data) represents one administration of the SF-36v2[®]. If a patient completed the SF-36v2[®] more than once, s/he should contribute one case in the data for each time s/he took the survey.

Please do not arrange the data so that one case represents several SF-36v2[®] administrations on one patient, one question on the SF-36v2[®], or any other unit of analysis.

- **Variables**

The variables in your dataset should have the names and values shown in Table 1. If possible, arrange them in this order at the start of the dataset. All variables listed here are **required** in your dataset, except for pid and timepoint, which are required only if you have purchased scoring for multiple timepoints.

Your dataset can also include any number of other variables. For example, you may have covariates you'd like to use in future analyses. Feel free to leave them in; we'll retain them and return them back to you. If extraneous variables are left in your dataset please ensure that none of these variables could be used to identify the patient in any way.

Table 1. Variables in submitted datasets

Variable name	Description	Variable type*
recordid	A unique identifier for each case (each administration of the SF-36v2 [®])	Numeric, any length
pid	A unique identifier for each subject, often called “Patient ID.” This variable is required only if you have purchased multiple timepoints	Numeric, any length
timepoint	The visit at which the survey was administered. This variable may take values such as (1,2,3...), (Baseline, Visit2, Vist4...) or (A,B,C...), etc. This variable is required only if you have purchased multiple timepoints	String or numeric, any length
form	Name of the form. This has the same value for all cases: SF-36v2 [®] . We use this as a double-check that the data is what we are expecting	String, length at least seven
recall	Recall period of the form, such as (1 week, 4 weeks),(one, four) or (1,4)	String or numeric, any length
GH01	Form question 1	All form questions: numeric, any length
HT	Form question 2	
PF01	Form question 3a	
PF02	Form question 3b	
PF03	Form question 3c	
PF04	Form question 3d	
PF05	Form question 3e	
PF06	Form question 3f	
PF07	Form question 3g	
PF08	Form question 3h	
PF09	Form question 3i	
PF10	Form question 3j	
RP01	Form question 4a	
RP02	Form question 4b	
RP03	Form question 4c	
RP04	Form question 4d	
RE01	Form question 5a	
RE02	Form question 5b	
RE03	Form question 5c	
SF01	Form question 6	
BP01	Form question 7	
BP02	Form question 8	
VT01	Form question 9a	
MH01	Form question 9b	
MH02	Form question 9c	
MH03	Form question 9d	
VT02	Form question 9e	
MH04	Form question 9f	
VT03	Form question 9g	
MH05	Form question 9h	
VT04	Form question 9i	
SF02	Form question 10	
GH02	Form question 11a	
GH03	Form question 11b	
GH04	Form question 11c	
GH05	Form question 11d	

* If applicable. See the “Acceptable file formats” section on the next page.

- **Acceptable file formats**

Our scoring software imports and exports files in .csv format (comma separated values), so we prefer your data in .csv if possible. All common statistical programs such as Excel, SAS, Stata and SPSS can export data in .csv format. Since .csv files do not distinguish between string and numeric variables, they need not follow the variable type requirements above.

In all, we can accept files with these formats:

- .csv Comma separated values, *preferred format*
- .xls or .xlsx Microsoft Excel format
- .sas7bdat SAS data format
- .xpt SAS data transport file (if using this format, please let us know via email whether you used a PROC COPY or DATA XPORT statement to create the dataset)

- **Missing values**

If your file is in .csv format, missing values must be represented by a period rather than a zero, space, null value, or other character. For any other format, missing values native to that platform are fine.

Please consider the distribution of missing values in your dataset before you submit the data:

- *Cases that are missing a response on a few items:*

We recommend keeping these cases in your data. Using missing data estimation techniques, our scoring software is able to use these cases as input to the scale scores and component scores, even when some items are missing.

- *Cases that are missing on all or most items:*

Keeping or dropping these cases before scoring is up to you. They will not affect the calculation of SF-36v2[®] scores, but they will affect the estimation of data quality, increasing the number of missing cases. Keep these cases if you have a special reason for including them -- they may be part of a carefully drawn sample, for example, and you want to stay aware of how many in your sample did not complete the SF-36v2[®]. Otherwise, dropping them is fine.

3. Transferring your data to us

We use the web- and email-based Accellion system to transfer data securely. For more information please refer to the separate Accellion instructions sent with these guidelines.

4. What you will receive back

- **Scored dataset**

The dataset you receive back from us will contain all the SF-36v2[®] scale and component score variables, with the names and labels as shown in Table 2. It will also retain any extra variables you had that weren't specifically required by us.

We will return the dataset in the same format in which it was received. For example, if you send a .csv file, we will return a .csv file.

Table 2. SF-36v2[®] derived variables that will be added to your data

Variable name	Variable label	Notes
PF	Physical functioning scale: 0-100 score	These are the raw 0-100 scores
RP	Role physical scale: 0-100 score	
BP	Bodily pain scale: 0-100 score	
GH	General health scale: 0-100 score	
VT	Vitality scale: 0-100 score	
SF	Social functioning scale: 0-100 score	
RE	Role emotional scale: 0-100 score	
MH	Mental health scale: 0-100 score	
PF_NBS	Physical functioning scale: norm-based score	These are the scores normed to the US population to have a mean of 50 and standard deviation of 10
RP_NBS	Role physical scale: norm-based score	
BP_NBS	Bodily pain scale: norm-based score	
GH_NBS	General health scale: norm-based score	
VT_NBS	Vitality scale: norm-based score	
SF_NBS	Social functioning scale: norm-based score	
RE_NBS	Role emotional scale: norm-based score	
MH_NBS	Mental health scale: norm-based score	
PCS	Physical component summary	
MCS	Mental component summary	
SF6D_R2	SF-6D health utility index score	See descriptions in Section 4 "What you will receive back"
MHE	Mental Health Enhanced score	
RCI	Response Consistency Index	

- **Additional variables**

The dataset will also include these variables:

- **SF6D Health Utility Index (SF6D_R2)**

By applying a scoring method that focuses on six health domains covered by the SF-36v2[®], we provide a preference-based health state utility index called SF-6D. This index can be used to conduct cost-utility and cost-benefit analyses of the relative benefits gained from treatment in terms of quality adjusted life years. The R2 in the variable name SF6D_R2 refers to the method of calculation of the SF6D.

- **Mental Health Enhanced score (MHE)**

Drawing on information from the five mental health items in the SF-36v2[®], we provide this enhanced mental health score, which is equivalent to standard depression inventories and can be used as a depression screening tool. The MHEnhanced score ranges from 0 (no depression) to 63 (severe depression).

- **Response Consistency Index (RCI)**

The RCI which evaluates the logical consistency of 15 pairs of responses to individual survey items. For example, a respondent's indication that they can "walk more than a mile" and at the same time cannot "walk one block," would be an inconsistent pair. The RCI counts the number of these inconsistencies: if all 15 response pairs are consistent, the RCI is zero, and if all are inconsistent the RCI is 15.

- **Data Quality Evaluation (DQE) report**

This written report contains:

- a. A general discussion of data quality issues and scaling assumptions of the SF-36v2[®]
- b. Information on your missing data rates and the amount of missing data estimation we performed
- c. Results of several scaling tests showing whether your data are consistent with the measurement model developed for the SF-36v2[®]
- d. Individualized interpretation of your data quality by an Optum analyst

If you have purchased multiple timepoints, you will receive a separate DQE for each timepoint.

5. Scoring timeline

Below is the expected timeline for your scoring project, assuming that the data we receive from you closely follows the specifications above. If not, some delays may result.

All days refer to business days.

- a. The day we receive the data is day one.
- b. By the close of day two, we will have checked for any major problems with formatting, structure, missing variables, etc. We will contact you only if there is an issue. If you don't hear from us on day two, you can assume that your data are fine and will be scored.
- c. We return the scored data and DQE report to you via Accellion by close of business on day four.

6. Follow-up

Please contact Scoring Help at scoringhelp@qualitymetric.com if you would like to schedule a call to review the DQE report, ask questions, or inquire about follow-up analytic work or other PRO projects our analysts and scientists can assist you with.