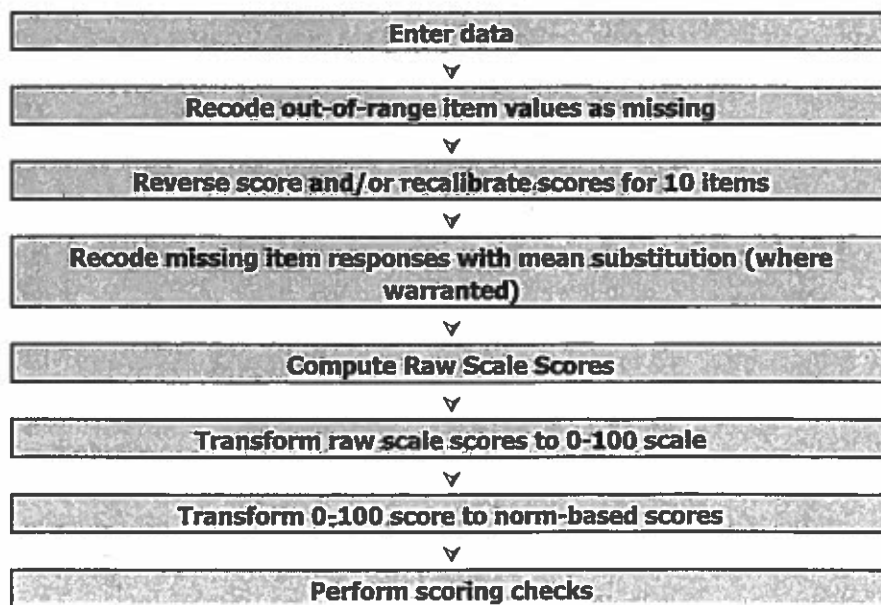# Chapter 6 SCORING SF-36 SCALES

This chapter provides scoring instructions for the eight multi-item scales and for the reported health transition item included in the SF-36 Health Survey, Version 2.0. General scoring information and steps for data entry and scoring that are common to all items are discussed first (see Figure 6.1). Next, formulas for item aggregation and transformation of scores to a 0-100 scale are presented. Following the 0-100 transformation is the norm-based scoring of each scale to have a mean of 50 and a standard deviation of 10 in the 1998 general U.S. population. Norm-based scoring methods of scales from standard (4-week recall) and acute (1 week recall) forms are presented separately. Finally, formal checks for errors in scoring are explained. A test dataset is available upon request (info@qmetric.com) to confirm successful reproduction of the scoring methods outlined in this chapter.

Importance of Standardization

As with all standardized tests, standardization of content and scoring is what makes interpretation of SF-36 scales possible. The content of the SF-36 form and the scoring algorithms were selected and standardized following careful study of many options. The algorithms described in this chapter were chosen to: be as simple as possible. satisfy the assumptions of the methods used to construct SF-36 scales, and to achieve the nearly perfect linear association with the original SF-36 scores, necessary to preserve their original interpretations.

Figure 6.1   Flow Chart for Scoring SF-36 Scales

Enter data

∨

Recode out-of-range item values as missing

∨

Reverse score and/or recalibrate scores for 10 items

∨

Recode missing item responses with mean substitution (where warranted)

∨

Compute Raw Scale Scores

∨

Transform raw scale scores to 0-100 scale

∨

Transform 0-100 score to norm-based scores

∨

Perform scoring checks

Changes in the content of the survey or in the scoring method may compromise the reliability and validity of scores. Changes are also likely to bias scores sufficiently to invalidate normative comparisons and to prevent comparisons of results across studies.

There are at least two good reasons to adhere to the standards of content and scoring described in this manual. First, they are most likely to produce scores with the same reliability and validity as those previously reported for SF-36v2 scales. Second, comparisons of results across studies are made possible to the benefit of all who use these content and scoring standards. Third, differences in scores will have the same interpretations.

Prior to using the SF-36v2 scoring rules, it is essential to verify that the questionnaires being scored, including the questions asked (item stems), response choices and numbers assigned to response choices at the time of data entry, have been reproduced exactly. The scoring rules in this chapter are appropriate for the standard SF-36 survey questions, response choices, and numbers assigned to response choices as reproduced in the *Appendices A-D*.

General Scoring Information

SF-36v2 items and scales are scored so that a higher score indicates a better health state. For example, functioning scales are scored so that a high score indicates better functioning and the pain scale is scored so that a high score indicates freedom from pain. After data entry, items and scales are scored in four steps:

1. Item recoding for the 10 items that require recoding
2. Computing raw scale scores by summing across items in the same scale (raw scale scores)
3. Transforming raw scale scores to a 0 - 100 scale (transformed scale scores)
4. Transforming 0-100 scale scores to have a mean of 50 and standard deviation of 10 in the general U.S. population (norm-based scale scores).

Note that for both standard and acute SF-36v2 forms, steps 1, 2 and 3 are the same for scoring the scales. However, step 4 differs for standard and acute forms. The transformation of 0-100 scales to norm-based scores requires a different set of means and standard deviations to standardize the scales from standard and acute forms.

Scoring & Data Quality Analysis Service

QualityMetric offers a scoring and data quality analysis service for both Version 1 and Version 2 of the SF-36 and SF-12. Information on this service can be found on the Internet at www.qmetric.com, or by sending an email to scoring@qmetric.com or by calling 401-334-8800.

Data Entry

The SF-36v2 item responses should be data entered as coded in the questionnaire. It is important to note that although the numbers printed along with the response choices should be data entered, they may not be the numbers ultimately assigned to those responses when SF-36 scales are scored.

In most cases, this means that the precoded number that is circled or marked by the respondent should be entered. However, sometimes it is not clear what number should be data-entered. Suggested rules for handling some of the more common coding problems are:

1. If a respondent marks two responses that are adjacent to each other, randomly pick one and enter that number.
2. If a respondent marks two responses for an item and they are not adjacent to each other, code that item "missing".
3. If a respondent marks three or more responses for an item, code that item "missing".

Several licensed vendors have developed scanning forms for use with the SF-36v2 in both standard and acute formats. Optical scanning generally reduces the time required to process questionnaires, but may involve greater initial investment in form design. Some scanning forms may require special processing equipment; however, this method may be cost-effective, especially if the SF-36 is being administered frequently or to a large sample. Other mechanisms for capturing data include interactive-voice response (IVR) systems, computer touch screen administration, and Internet based systems. For a listing of licensed vendors, please visit the QualityMetric Web site (www.qmetric.com).

Tables 6.1 through 6.9 present scoring information for the items used in each of the eight SF-36 health scales and the self-reported health transition item. Each table presents the verbatim content of each question, response choices, and both the precoded values printed in the questionnaire and final values for scoring each item. Item numbers in Tables 6.1 through 6.9 correspond to those on the Standard and Acute SF-36 forms (reproduced in the Appendix).

Table 6.1 Physical Functioning: Verbatim Items and Scoring Information

Question No.	Variable Label	VERBATIM ITEMS
3a.	PF01	Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports
3b.	PF02	Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf
3c.	PF03	Lifting or carrying groceries
3d.	PF04	Climbing several flights of stairs
3e.	PF05	Climbing one flight of stairs
3f.	PF06	Bending, kneeling, or stooping
3g.	PF07	Walking more than a mile
3h.	PF08	Walking several hundred yards
3i.	PF09	Walking one hundred yards
3j.	PF10	Bathing or dressing yourself

PRECODED AND FINAL VALUES FOR ITEMS 3A - 3J		
Response Choices	Precoded Item Value	Final Item Value
Yes, limited a lot	1	1
Yes, limited a little	2	2
No, not limited at all	3	3

Scale Scoring

Compute the simple algebraic sum of the final item scores as shown in Table 6.11. See text for handling of missing item responses. This scale is scored so that a high score indicates better physical functioning.

Note: Precoded values are as shown on the appended form.
This scale does not require recoding of items prior to computation of the scale score.

Table 6.2 Role-Physical: Verbatim Items and Scoring Information

Question No.	Variable Label	VERBATIM ITEMS
4a.	RP01	Cut down the **amount of time** you spent on work or other activities
4b.	RP02	**Accomplished less** than you would like
4c.	RP03	Were limited in the **kind** of work or other activities
4d.	RP04	Had **difficulty** performing the work or other activities (for example, it took extra effort)

PRECODED AND FINAL VALUES FOR ITEMS 4A - 4D		
Response Choices	**Precoded Item Value**	**Final Item Value**
All of the time	1	1
Most of the time	2	2
Some of the time	3	3
A little of the time	4	4
None of the time	5	5

Scale Scoring

Compute the simple algebraic sum of the final item scores as shown in Table 6.11. See text for handling of missing item responses. This scale is scored so that a high score indicates better Role-Physical functioning.

Note: Precoded values are as shown on the appended form.
This scale does not require recoding of items prior to computation of the scale score.

Table 6.3 Bodily Pain: verbatim Items and Scoring Information   Bodily Pain: Verbatim Items and Scoring Information

Question No.	Variable Label	VERBATIM ITEMS
7.	BP01	How much bodily pain have you had during the past 4 weeks?
8.	BP02	During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

PRECODED AND FINAL VALUES FOR ITEM 7		
Response Choices	Precoded Item Value	Final Item Value
None	1	6.0
Very mild	2	5.4
Mild	3	4.2
Moderate	4	3.1
Severe	5	2.2
Very severe	6	1.0

SCORING FOR ITEM 8 IF BOTH ITEMS 7 AND 8 ARE ANSWERED			
Response Choices	If Item 8 Precoded Item Value	&Item 7 Precoded Item Value	Then Final Item Value
Not at all	1	1	6
Not at all	1	2 through 6	5
A little bit	2	1 through 6	4
Moderately	3	1 through 6	3
Quite a bit	4	1 through 6	2
Extremely	5	1 through 6	1

SCORING FOR ITEM 8 IF ITEM 7 IS NOT ANSWERED		
Response Choices	Precoded Item Value	Then Final Item Value
Not at all	1	6.0
A little bit	2	4.75
Moderately	3	3.5
Quite a bit	4	2.25
Extremely	5	1.0

Scale Scoring

Compute the simple algebraic sum of the final item scores as shown in Table 6.11. See text for handling of missing item responses. This scale is scored so that a high score indicates lack of bodily pain.

Note: Precoded values are as shown on the appended form.
This scale requires recoding of both items prior to computation of the scale score.

Table 6.4 General Health: Verbatim Items and Scoring Information

Question No.	Variable Label	VERBATIM ITEMS
1.	GH01	In general, would you say your health is:
11a.	GH02	I seem to get sick a little easier than other people
11b	GH03	I am as healthy as anybody I know
11c.	GH04	I expect my health to get worse
11d.	GH05	My health is excellent

PRECODED AND FINAL VALUES FOR ITEMS 1 & 11A - 11D			
Item 1	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	Excellent	1	5.0
	Very good	2	4.4
	Good	3	3.4
	Fair	4	2.0
	Poor	5	1.0
Items 11a & 11c	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	Definitely True	1	1
	Mostly True	2	2
	Don't Know	3	3
	Mostly False	4	4
	Definitely False	5	5
Items 11b & 11d	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	Definitely True	1	5
	Mostly True	2	4
	Don't Know	3	3
	Mostly False	4	2
	Definitely False	5	1

Scale Scoring

Compute the simple algebraic sum of the final item scores as shown in Table 6.11. See text for handling of missing item responses. This scale is scored so that a high score indicates better general health perceptions.

Note: Precoded values are as shown on the appended form.
This scale requires recoding of three items prior to computation of the scale score.

Table 6.5 Vitality: Verbatim Items and Scoring Information Vitality: Verbatim Items and Scoring Information

Question No.	Variable Label	VERBATIM ITEMS
9a.	VT01	Did you feel full of life?
9e.	VT02	Did you have a lot of energy?
9g.	VT03	Did you feel worn out?
9i.	VT04	Did you feel tired?

PRECODED AND FINAL VALUES FOR ITEMS 9A, 9E, 9G, & 9I			
Item 9a & 9e	Response Choices	Precoded Item Value	Final Item Value
	All of the time	1	5
	Most of the time	2	4
	Some of the time	3	3
	A little of the time	4	2
	None of the time	5	1
Items 9g & 9i	Response Choices	Precoded Item Value	Final Item Value
	All of the time	1	1
	Most of the time	2	2
	Some of the time	3	3
	A little of the time	4	4
	None of the time	5	5

Scale Scoring

Compute the simple algebraic sum of the final item scores as shown in Table 6.11. See text for handling of missing item responses. This scale is scored so that a high score indicates more vitality.

Note: Precoded values are as shown on the appended form.
This scale requires recoding of two items prior to computation of the scale score.

Table 6.6   Social Functioning: Verbatim Items and Scoring Information

Question No.	Variable Label	VERBATIM ITEMS
6.	SF01	During the **past 4 weeks**, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?
10.	SF02	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.)?

PRECODED AND FINAL VALUES FOR ITEMS 6 & 10			
Item 6	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	Not at all	1	5
	Slightly	2	4
	Moderately	3	3
	Quite a bit	4	2
	Extremely	5	1
Items 10	**Response Choices**	**Precoded Item Value**	**Final Item Value**
	All of the time	1	1
	Most of the time	2	2
	Some of the time	3	3
	A little of the time	4	4
	None of the time	5	5

Scale Scoring

Compute the simple algebraic sum of the final item scores as shown in Table 6.11. See text for handling of missing item responses. This scale is scored so that a high score indicates better social functioning.

Note: Precoded values are as shown on the appended form.
This scale requires recoding of one item prior to computation of the scale score.

Table 6.7 Role-Emotional: Verbatim Items and Scoring Information

Question No.	Variable Label	VERBATIM ITEMS
5a.	RE01	Cut down the **amount of time** you spent on work or other activities
5b.	RE02	**Accomplished less** than you would like
5c.	RE03	Did work or other activities less **carefully** than usual

PRECODED AND FINAL VALUES FOR ITEMS 5A - 5C			
Items 5a – 5c	Response Choices	Precoded Item Value	Final Item Value
	All of the time	1	1
	Most of the time	2	2
	Some of the time	3	3
	A little of the time	4	4
	None of the time	5	5

Scale Scoring

Compute the simple algebraic sum of the final item scores as shown in Table 6.11. See text for handling of missing item responses. This scale is scored so that a high score indicates better Role-Emotional functioning.

Note: Precoded values are as shown on the appended form.
This scale does not require recoding of items prior to computation of the scale score.

Table 6.8 Mental Health: Verbatim Items and Scoring Information

Question No.	Variable Label	VERBATIM ITEMS
9b.	MH01	Have you been very nervous?
9c.	MH02	Have you felt so down in the dumps that nothing could cheer you up?
9d.	MH03	Have you felt calm and peaceful?
9f.	MH04	Have you felt downhearted and depressed?
9h	MH05	Have you been happy?

PRECODED AND FINAL VALUES FOR ITEMS 9B, 9C, 9D, 9F, & 9H			
Items 9b, 9c & 9f	Response Choices	Precoded Item Value	Final Item Value
	All of the time	1	1
	Most of the time	2	2
	Some of the time	3	3
	A little of the time	4	4
	None of the time	5	5
Items 9d & 9h	Response Choices	Precoded Item Value	Final Item Value
	All of the time	1	5
	Most of the time	2	4
	Some of the time	3	3
	A little of the time	4	2
	None of the time	5	1

Scale Scoring

Compute the simple algebraic sum of the final item scores as shown in Table 6.11. See text for handling of missing item responses. This scale is scored so that a high score indicates better mental health.

Note: Precoded values are as shown on the appended form.
This scale requires recoding of two items prior to computation of the scale score.

Table 6.9 Reported Health Transition: Verbatim Items and Scoring Information

Question No.	Variable Label	VERBATIM ITEMS
2.	HT	**Compared to one year ago**, how would you rate your health in general now?

PRECODED AND FINAL VALUES FOR ITEM 2			
Item 2	Response Choices	Precoded Item Value	Final Item Value
	Much better now than one year ago	1	1
	Somewhat better now than one year ago	2	2
	About the same as one year ago	3	3
	Somewhat worse now than one year ago	4	4
	Much worse now that one year ago	5	5

Note: *Precoded values are as shown on the appended form.*

Item Recoding

The next stage after data entry is the recoding of response choices as shown in Tables 6.1 through 6.9. Item recoding is the process of deriving the item values that will be used to calculate the scale scores. Several steps are included in this process: (1) change out-of-range values to missing, (2) recode values for 10 items, and (3) substitute person-specific estimates for missing items.

Out of Range Values

All 36 items should be checked for out-of-range values prior to assigning the final item values. Out-of-range values are those that are lower than an item's precoded minimum value, or higher than an item's precoded maximum value (see Tables 6.1 through 6.9). Out-of-range values are usually caused by data-entry errors and, if possible, should be changed to the correct response through verification with the original questionnaire. If the questionnaire is not available, all out-of-range values should be recoded as missing data.

Recode Values for Ten Items

Ten items are reverse scored. Reverse scoring of items is done to ensure that a higher item value indicates better health on all SF-36v2 items scales. SF-36 items that need to be reverse scored are worded so that a higher precoded item value indicates a poorer health state.

Item Recalibration

For 34 of the SF-36v2 items, research to date offers good support for the assumption of a linear relationship between item scores and the underlying health concept defined by their scales. However, empirical work has shown that two items require recalibration to satisfy this important scaling assumption. These items are in two different SF-36 scales: the General Health (GH) scale and the Bodily Pain (BP) scale.

General Health Rating Item

The "Very Good" and "Good" responses to Item 1 (GH1) are recalibrated to achieve a better linear fit with the general health evaluation concept measured by the GH scale. Empirical studies during the Health Insurance Experiment (HIE) were among the first to document that the intervals between response choices for this item are not equal (Davies & Ware, 1981). Subsequent studies of Item 1 (GH1), using both the Thurstone Method of Equal-Appearing Intervals (Thurstone & Chave, 1929) and other empirical methods, have also consistently shown that the interval between "Excellent" and "Very Good" is about half the size of the interval between "Fair" and "Good" (Ware, Nelson et al., 1992). These results have been confirmed in studies of SF-36 translations from 10 countries participating in the International Quality of Life Assessment (IQOLA) project. Finally, in all studies we are aware of to date, mean values for a criterion general health scale for respondents who choose each of the five levels defined by Item 1 (GH1) depart significantly from linearity.

Results from the Medical Outcome Study (MOS) that served as the basis for the recommended recalibration of Item 1 (GH1) are summarized in Table 6.10. As shown in Table 6.10 and discussed elsewhere (Ware, Nelson, et al., 1992), the mean criterion scores were remarkably similar for those who chose the same category of Item 1 across the screening (N=18,573) and longitudinal (N=3,054) samples. Intervals between adjacent response categories were unequal, as observed in the HIE (Davies & Ware, 1981). For these reasons, item scale values are transformed as shown in Table 6.10 using specific results from the screening sample. The result is a very high 0.70 correlation with the sum of the other four items in the GH scale.

Table 6.10 Mean Current Health Scores for Respondents
Choosing Each Level of SF-36® Item 1 (GH1)

Response to Item 1	Mean Current Health		Recommended Scoring	
	Screening Sample (N=18,573)	Baseline Sample (N=3,054)	1-5 Scale	0-100 Scale
Excellent	87.9	86.9	5.0	100
Very good	75.5	75.4	4.4	84
Good	57.6	55.9	3.4	61
Fair	30.0	30.6	2.0	25
Poor	10.8	10.8	1.0	0

Note. Adapted from: Ware, Nelson et al. (1992). Preliminary tests of a 6-item general health survey: A patient application. In A.L. Stewart & J.E. Ware (Eds.), Measuring functioning and well-being: The Medical Outcomes Study approach (p. 299). Durham, NC: Duke University Press.

Bodily Pain Items

The scoring rules recommended for the Bodily Pain (BP) scale were based on three considerations: (1) the items offer both different numbers and different content of response choices, (2) administration of Item 8 (BP2) depended on the response to an item like Item 7 (BP1) in the MOS, and (3) empirical studies indicate that recalibration of Item 7 (BP1) is necessary to achieve a linear fit with the scale score and with other measures of bodily pain.

As shown in Table 6.3, the two bodily pain items offer an unequal number of response choices (six for Item 7 and five for Item 8). As a result, their variances are not equal, as required for a summated rating scale. Further, in all MOS studies published to date, Item 8 was administered (following a skip pattern) only to those respondents reporting at least some pain. Although the MOS skip pattern has been dropped to make the SF-36v2 easier to administer, the dependence between responses must be taken into account to compare results from new studies with published studies.

The recommended recoding of the first response choice for Item 8 (BP2) on the basis of the response to Item 7 (BP1) solves two problems. First, it converts Item 8 to a six-level item of roughly equal variance to Item 7. This is done by splitting those free of role interference due to pain into two different groups: (1) free of interference and free of pain (the best level); and (2) free of interference but with at least some pain (the next best level). Second, it approximates the dependence between the two items in MOS studies of reliability and validity to date (McHorney et al., 1992, 1993).

Davies and Ware (1981) reported that recalibration of the bodily pain severity rating was necessary to satisfy the equal interval assumption in studies during the HIE. MOS studies have confirmed that the relationship between Item 7 (BP1) and criterion measures of pain departs significantly from a linear association. Criterion pain measures used in these tests include visual analogue scales measuring pain severity and categorical ratings of pain frequency and duration. Final response values for Item 7 were derived from the mean values of a summary MOS criterion pain measure computed for respondents who chose each of the six levels defined by Item 7, using methods much like those illustrated in Table 6.10 for Item 1 (GH1).

How to Treat Missing Data

Sometimes respondents leave one or more questionnaire items in a scale blank, although this happens infrequently (1 to 2% or less) in most surveys. One important advantage of multi-item scales is that a scale score can be estimated even though responses to some items are missing. Using a scoring algorithm that estimates missing values, it is usually possible to derive scale scores for nearly all respondents across the eight SF-36v2 scales.

We recommend that a scale score be calculated if a respondent answered at least half of the items in a multi-item scale (or half plus one in the case of scales with an odd number of items).

The recommended algorithm substitutes a person-specific estimate for any missing item when the respondent answered at least 50 percent of the items in a scale. A psychometrically sound estimate is the average score, across completed items in the same scale, for that respondent (Ware, Davies-Avery, & Brook, 1980). For example, if a respondent leaves one item in the 5-item Mental Health scale blank, substitute the respondent's average score (across the four completed mental health items) for that one item. When estimating the respondent's average score, use the respondent's final item values, as defined in Tables 6.1 through 6.9. This step is easy to program using standard statistical software packages (e.g., SPSS, SAS). An example of SAS statistical software program code is available at QualityMetric, Inc. (info@qmetric.com).

At the time of this publication, new software was being evaluated for use in removing bias in estimates of scores for those having one or more missing SF-36 responses and to enable score estimation for virtually all respondents, regardless of the amount of missing data (see Kosinski, Bayliss, Bjorner and Ware 2000). For example, this software has made it possible to estimate summary scores for 88% of elderly respondents with missing data in the Medicare Managed Care Health Outcomes Survey (HOS). Scoring services using this software will be offered on the Internet at www.sf-36.com.

Computing Raw Scale Scores

After item recoding, including handling of missing data, a raw score is computed for each scale. This score is the simple algebraic sum of responses for all items in that scale, as shown in Table 6.11. For example, the raw scale score for the Role-Physical Scale is the sum of the scores for Items 4a, 4b, 4c, and 4d. Use recoded item values and imputed values where applicable. Generally, we recommend that if the respondent answers at least 50% of the items in a multi-item scale, the score should be calculated. If the respondent did not answer at least 50% of the items, the score for that scale should be set to missing. Some prefer a more conservative approach for the scales with only two items, and set those scales to missing unless both items are completed.

This simple scoring method is possible because items in the same scale have roughly equivalent relationships to the underlying health concept being measured and no item is used in more than one scale. Thus, it is not necessary to standardize or weight items. These assumptions have been extensively tested and verified for both Versions 1.0 and 2.0 (McHorney et al., 1994; Ware 1993; see chapter 4).

Transformation of Scale Scores

The next step involves transforming each raw scale score to a 0 - 100 scale using the formula shown below. Table 6.11 provides the information necessary to apply this formula to each scale.

$$\text{Transformed Scale} = \left[\frac{(\text{Actual raw score} - \text{lowest } \textit{possible} \text{ raw score})}{\textit{Possible} \text{ raw score range}} \right] * 100$$

This transformation converts the lowest and highest possible scores to 0 and 100, respectively. Scores between these values represent the percentage of the total possible score achieved. Raw and transformed scale scores are not calculated for the Reported Health Transition item. We recommend treating responses to this item as ordinal level data and analyzing the percentage of respondents who select each response choice or using the estimates of measured change (observed changes in SF-36v2 scale scores) reported for each response category.

Table 6.11   Formulas for Scoring and Transforming Scales

Scale	Sum Final Item Values (after recoding items as in Table 1.2-1.9)	Lowest and highest possible raw scores	Possible raw score range
Physical Functioning	3a+3b+3c+3d+3e+3f+3g+3h+3i+3j	10, 30	20
Role-Physical	4a+4c+4c+4d	4, 20	16
Bodily Pain	7+8	2, 12	10
General Health	1+11a+11b+11c+11d	5, 25	20
Vitality	9a+9e+9g+9I	4, 20	16
Social Functioning	6+10	2, 10	8
Role-Emotional	5a+5b+5c	3, 15	12
Mental Health	9b+9c+9d+9f+9h	5, 25	20

Formula and example for transformation of raw scale scores to 0-100 scale scores

$$\text{Transformed Scale} = \left\lceil \frac{(\text{Actual raw score - lowest } \textit{possible} \text{ raw score})}{\textit{Possible} \text{ raw score range}} \right\rceil * 100$$

Example: A Physical Functioning raw score of 21 would be transformed as follows:

$$\left\lceil \frac{(21 - 10)}{20} \right\rceil * 100 = 55$$

Where lowest possible score = 10 and possible raw score range = 20

Norm-Based Scoring of Scale Scores, Standard Form (4-Week Recall)

The next step involves the norm-based scoring of each 0 - 100 scale score using the formulas shown below. Table 6.12 provides the information necessary to apply these formulas to each SF-36v2 scale from the standard form (4 week recall). The means and standard deviations (Table 6.12) used in norm-based scoring come from the 1998 general U.S. population. A linear z-score transformation is used so that all eight SF-36 scales have a mean of 50 and a standard deviation of 10 in the 1998 general U.S. population.

The advantage of the standardization and norm-based scoring of the 8 SF-36v2 scales is that results for one scale can be meaningfully compared with the other scales and their scores have a direct interpretation in relation to the distribution of scores in the 1998 general U.S. population. Specifically, all scores above or below 50 are above or below the average, respectively, in the 1998 general U.S. population. Because the standard deviation is 10 for all 8 scales, each one point difference or change in scores also has a direct interpretation. A one point difference or change is one-tenth of a standard deviation unit or an effect size of 0.10. Lastly, norm-based scoring provides the basis for comparing scale scores across Version 1.0 and Version 2.0 standard forms.

The first step in norm-based scoring consists of standardizing each SF-36v2 scale using a z-score transformation. A z-score for each scale is computed by subtracting the 1998 general U.S. population mean (Table 6.12) for each SF-36 scale and dividing the difference by the corresponding scale standard deviation (Table 6.12) from the 1998 general U.S. population. Formulas are listed below.

Step 1. Formulas for z-score standardization of SF-36v2 scales (Standard Form)

$$PF\_Z = (PF - 83.29094) / 23.75883$$
$$RP\_Z = (RP - 82.50964) / 25.52028$$
$$BP\_Z = (BP - 71.32527) / 23.66224$$
$$GH\_Z = (GH - 70.84570) / 20.97821$$
$$VT\_Z = (VT - 58.31411) / 20.01923$$
$$SF\_Z = (SF - 84.30250) / 22.91921$$
$$RE\_Z = (RE - 87.39733) / 21.43778$$
$$MH\_Z = (MH - 74.98685) / 17.75604$$

Means and standard deviations are from Table 6.12.

Table 6.12   1998 General U.S. Population Means and Standard
Deviations Used to Derive SF-36v2™ z-scores
(Standard form)

SF-36® Scale	Mean	Standard Deviation
PF	83.29094	23.75883
RP	82.50964	25.52028
BP	71.32527	23.66224
GH	70.84570	20.97821
VT	58.31411	20.01923
SF	84.30250	22.91921
RE	87.39733	21.43778
MH	74.98685	17.75604

Step 2. Norm-based Transformation of SF-36v2™ Z-Scores

The second step involves transforming each SF-36v2 z-score to the norm-based (50, 10) scoring. This is accomplished by multiplying each z-score from Step 1 by 10 and adding the resulting product to 50. Formulas are listed below.

Step 2.   Norm-based transformation of SF-36v2™ z-scores (Standard Form):

Norm-Based PF: $PF = 50 + (PF\_Z * 10)$

Norm-Based RP: $RP = 50 + (RP\_Z * 10)$

Norm-Based BP: $BP = 50 + (BP\_Z * 10)$

Norm-Based GH: $GH = 50 + (GH\_Z * 10)$

Norm-Based VT: $VT = 50 + (VT\_Z * 10)$

Norm-Based SF: $SF = 50 + (SF\_Z * 10)$

Norm-Based RE: $RE = 50 + (RE\_Z * 10)$

Norm-Based MH: $MH = 50 + (MH\_Z * 10)$

Scoring Checks

Because of errors in reproducing a form, entering data, programming or processing, which can lead to inaccurate scale scores, we strongly recommend formal scoring checks prior to using the scales. Any discrepancies observed during the following checks should be investigated for scoring errors:

1. Calculate SF-36v2 scale scores by hand for several respondents and compare the results to those produced by your scale scoring computer software.
2. After items have been coded into their final item values, inspect the frequency distributions for the items to verify that only the final item values shown in Tables 6.1 through 6.9 are observed. Discrepancies should be limited to respondents with values estimated for missing data as described above.

3. After items have been recoded and scale scores have been computed, inspect the correlation between each scale and its component items to verify that all correlations are positive in direction and substantial in magnitude (0.30 or higher).
4. Check correlation's between the General Health scale and the other seven scales to verify that all are positive; with rare exceptions they should also be substantial in magnitude (0.30 or higher).
5. For those familiar with principal factor or components analysis, inspect correlations between the eight scales and the first unrotated factor or component extracted from the correlation's among those scales. Regardless of extraction method, these correlations should be positive and substantial in magnitude (0.30 or higher).

Norm-Based Scoring of Scale Scores, Acute Form (1-week Recall)

After scoring the acute version SF-36v2 scales on a 0-100 scale, the next step involves the norm-based scoring (50/10) using the formulas shown below. Table 6.13 provides the information necessary to apply these formulas to each SF-36 scale from the Acute form (1 week recall). The means and standard deviations (Table 6.13) used in norm-based scoring come from the 1998 general U.S. population. A linear z-score transformation is used so that all eight SF-36 scales have a mean of 50 and a standard deviation of 10 in the 1998 general U.S. population.

The advantages of the norm-based scoring of the standard recall form above also apply to the acute version. The methods for transforming 0-100 scores on the acute scales to norm-based scores are also similar to those used for the standard form. The first step in norm-based scoring consists of standardizing each SF-36v2 scale using a z-score transformation. A z-score for each scale is computed by subtracting the 1998 general U.S. population mean (Table 6.13) for each acute form SF-36 scale and dividing the difference by the corresponding scale standard deviation (Table 6.13) from the 1998 general U.S. population. Formulas are listed below.

Step 1.   **Formulas for z-score standardization of SF-36v2™ scales (Acute Form)**

$$PF\_Z = (PF - 82.62455) / 24.43176$$
$$RP\_Z = (RP - 82.65109) / 26.19282$$
$$BP\_Z = (BP - 73.86999) / 24.00884$$
$$GH\_Z = (GH - 70.78372) / 21.28902$$
$$VT\_Z = (VT - 58.41968) / 20.87823$$
$$SF\_Z = (SF - 85.11568) / 23.24464$$
$$RE\_Z = (RE - 87.50009) / 22.01216$$
$$MH\_Z = (MH - 75.76034) / 18.04746$$

Means and standard deviations are from Table 6.13.

Table 6.13   1998 General U.S. Population Means and Standard Deviations Used to Derive SF-36 Z-Scores (Acute Form)

SF-36® Scale	Mean	SD
PF	82.62455	24.43176
RP	82.65109	26.19282
BP	73.86999	24.00884
GH	70.78372	21.28902
VT	58.41968	20.87823
SF	85.11568	23.24464
RE	87.50009	22.01216
MH	75.76034	18.04746

Step 2. Norm-based Transformation of SF-36v2™ Z-Scores (Acute Form)

The second step involves transforming each SF-36v2 z-score from Step 1 to the norm-based (50, 10) scoring. This is accomplished by multiplying each z-score by 10 and adding the resulting product to 50. Formulas are listed below.

Step 2.   Norm-based transformation of SF-36v2™ z-scores (Acute Form):

Norm-Based PF: $PF = 50 + (PF\_Z * 10)$

Norm-Based RP: $RP = 50 + (RP\_Z * 10)$

Norm-Based BP: $BP = 50 + (BP\_Z * 10)$

Norm-Based GH: $GH = 50 + (GH\_Z * 10)$

Norm-Based VT: $VT = 50 + (VT\_Z * 10)$

Norm-Based SF: $SF = 50 + (SF\_Z * 10)$

Norm-Based RE: $RE = 50 + (RE\_Z * 10)$

Norm-Based MH: $MH = 50 + (MH\_Z * 10)$

Scoring Checks

Because of errors in reproducing a form, entering data, programming or processing, which can lead to inaccurate scale scores, we strongly recommend formal scoring checks prior to using the scales. Any discrepancies observed during the following checks should be investigated for scoring errors:

1. Calculate SF-36v2 scale scores by hand for several respondents and compare the results to those produced by your scale scoring computer software.
2. After items have been coded into their final item values, inspect the frequency distributions for the items to verify that only the final item values shown in Tables 6.1 through 6.9 are observed. Discrepancies should be limited to respondents with values estimated for missing data as described above.
3. After items have been recoded and scale scores have been computed, inspect the correlation between each scale and its component items to verify that all correlations are positive in direction and substantial in magnitude (0.30 or higher).
4. Check correlation's between the General Health scale and the other

seven scales to verify that all are positive; with rare exceptions they should also be substantial in magnitude (0.30 or higher).

5. For those familiar with principal factor or components analysis, inspect correlations between the eight scales and the first unrotated factor or component extracted from the correlations among those scales. Regardless of extraction method, these correlations should be positive and substantial in magnitude (0.30 or higher).

Chapter 7   Scoring SF-36v2™ Physical and Mental Summary Measures

This chapter provides scoring instructions for the SF-36v2 physical (PCS) and mental (MCS) component summary measures. Scoring of the SF-36v2 PCS and MCS summary measures involves three steps. First, the eight SF-36 scales are standardized using means and standard deviations from the 1998 general U.S. population. Second, they are aggregated using weights (factor score coefficients) from the 1990 general U.S. population. These are the same weights as those used to score PCS and MCS from the SF-36 Version 1.0 (Ware et al., 1994). Finally aggregate PCS and MCS scores are standardized using a linear T-score transformation to have a mean of 50 and a standard deviation of 10, in the 1998 general U.S. population. Norm-based scoring of the standard (4-week recall) and acute (1 week recall) forms is presented separately.

General U.S. population statistics used in the standardization and in the aggregation of SF-36v2 scale scores are presented in Table 7.1 for the standard form and in Table 7.2 for the acute form. Note that for both standard and acute forms the same factor score coefficients are used to score PCS and MCS. To make sure that the original coefficients used in this scoring "recipe" have the same affect in 1998, as in 1990, 1998 variances are utilized. Detailed information including formulas for scale aggregation and transformation of scores are presented below. Formal checks using a test dataset available upon request (info@qmetric.com) can be performed to confirm the successful reproduction of PCS and MCS scales, as discussed later in the chapter. We strongly recommend these tests.

Importance of Standardization

As with the 1998 NBS scoring of SF-36v2 scales, which are aggregated to score the summary measures, standardization of the scoring of the PCS and MCS scales is vital to their interpretation. Any changes in scoring of the SF-36 scales or the algorithms for the summary measures may compromise their reliability and validity. Changes in scoring may also invalidate normative comparisons, based on the 1998 norms documented here.

The methods documented here achieve 1998 scores that are a nearly perfect transformation of the 1990 scores. For scales and summary measures, NBS based on 1998 norms only shifts the score distribution to better reflect the health of the US population in 1998. Otherwise, 1998 scores have the same interpretations as 1990 scores.

Norm-Based Scoring of PCS and MCS, Standard Form (4 Week Recall)

Standard form PCS and MCS scales are scored using norm-based methods. The means and standard deviations used in scoring come from the 1998 general U.S. population and the factor score coefficients come from the 1990 general U.S. population (Ware et al. 1994). A linear T-score transformation method is used so that both the PCS and MCS have a mean of 50 and a standard deviation of 10 in the 1998 general U.S. population.

The advantage of the standardization and norm-based scoring of the PCS and MCS is that results for one can be meaningfully compared with the other and their scores have a direct interpretation in relation to the distribution of scores in the general U.S. population. Specifically, all scores above and below 50 are above and below the average, respectively, in the 1998 general U.S. population. Because the standard deviation is 10 for both PCS and MCS measures, each one point difference in scores also has a direct interpretation. A one point difference is one-tenth of a standard deviation. (See Chapter 5 for an example of these advantages.)

Table 7.1   **1998 General U.S. Population Means, Standard Deviations and 1990 Factor Score Coefficients used to Derive PCS and MCS Scale Scores, Standard Form**

			Factor Score Coefficients	
SF-36® Scale	**Mean**[*]	**Standard Deviation**	**PCS**	**MCS**
PF	83.29094	23.75883	0.42402	-0.22999
RP	82.50964	25.52028	0.35119	-0.12329
BP	71.32527	23.66224	0.31754	-0.09731
GH	70.84570	20.97821	0.24954	-0.01571
VT	58.31411	20.01923	0.02877	0.23534
SF	84.30250	22.91921	-0.00753	0.26876
RE	87.39733	21.43778	-0.19206	0.43407
MH	74.98685	17.75604	-0.22069	0.48581

Note: The means and standard deviations for each SF-36v2™ are based on the 0-100 scoring.

It should be noted that mean scores for most of the scales vary substantially form those from the 1990 norms used to score Version 1. For those scales (PF, BP and GH) that have not been changed in any way from Version 1 to 2, several explanations are being explored (e.g., sampling bias in the 1990 study, downward temporal trends in physical health). Regardless of the explanations, these differences underscore the importance of using up-to-date norms and also the importance of using the same normative data when equating Versions 1 and 2 of the SF-36. 1998 norms for both make these comparisons possible.

Steps in Scoring

Following the scoring of the eight scales according to the standard SF-36v2 scoring algorithms (0-100 scale) explained earlier in chapter 6, PCS and MCS are scored in three steps as explained below:

Standardization of Scales (Z-Scores), Standard Form

The first step in computing PCS and MCS consists of standardizing each of the 8 SF-36 scales using a z-score transformation. This is the same as Step 1 used in the norm-based scoring of the 8 SF-36 scales explained in Chapter 6. A z-score for each scale is computed by subtracting the mean 0-100 general US population score (see Table 7.1) for each SF-36 scale and dividing the difference by the corresponding scale standard deviation. Note that the SF-36 scales scored on the 0-100 scale are used in step 1. Norm-based SF-36 scale scores are not used in this step. Formulas are listed below.

Step 1.　Formulas for z-score standardization of SF-36v2™scales, (Standard Form)

$$PF\_Z = (PF - 83.29094) / 23.75883$$
$$RP\_Z = (RP - 82.50964) / 25.52028$$
$$BP\_Z = (BP - 71.32527) / 23.66224$$
$$GH\_Z = (GH - 70.84570) / 20.97821$$
$$VT\_Z = (VT - 58.31411) / 20.01923$$
$$SF\_Z = (SF - 84.30250) / 22.91921$$
$$RE\_Z = (RE - 87.39733) / 21.43778$$
$$MH\_Z = (MH - 74.98685) / 17.75604$$

Means and standard deviations are from Table 7.1.

Aggregation of Scale Scores (Standard Form)

After a z-score has been computed for each SF-36v2 scale, the second step involves computation of aggregate scores for the physical and mental components using the physical and mental factor score coefficients from the 1990 general U.S. population as given in Table 7.1.

Computation of an aggregate physical component score consists of multiplying each SF-36v2 scale z-score by its respective physical factor score coefficient and summing the eight products, as shown below. Similarly, an aggregate mental component score is obtained by multiplying each SF-36 scale z-score by its respective mental factor score coefficient and summing the eight products.

Step 2.　Formulas for aggregating scales in estimating aggregate physical and mental component scores (Standard Form)

$$AGG\_PHYS = (PF\_Z * .42402) + (RP\_Z * .35119) + (BP\_Z * .31754) + (GH\_Z * .24954) + (VT\_Z * .02877) + (SF\_Z * -.00753) + (RE\_Z * -.19206) + (MH\_Z * -.22069)$$

$$AGG\_MENT = (PF\_Z * -.22999) + (RP\_Z * -.12329) + (BP\_Z * -.09731) + (GH\_Z * -.01571) + (VT\_Z * .23534) + (SF\_Z * .26876) + (RE\_Z * .43407) + (MH\_Z * .48581)$$

Transformation of Summary Scores

The third step involves transforming each component score to the norm-based (50, 10) scoring. This is accomplished by multiplying each aggregate component scale score by 10 and adding the resulting product to 50. Formulas are shown in step 3.

Step 3. Formulas for T-score transformation of component scores (Standard Form)

Transformed Physical (PCS) = 50 + (AGG_PHYS * 10)

Transformed Mental (MCS) = 50 + (AGG_MENT * 10)

Norm-Based Scoring of PCS and MCS, Acute Form

Acute form PCS and MCS scales are scored using norm-based methods. The means and standard deviations used in scoring come from the 1998 general U.S. population and the factor score coefficients come from the 1990 general U.S. population (Ware et al. 1994). A linear T-score transformation method is used so that both the PCS and MCS have a mean of 50 and a standard deviation of 10 in the 1998 general U.S. population.

The advantage of the standardization and norm-based scoring of the PCS and MCS is that results for one can be meaningfully compared with the other and their scores have a direct interpretation in relation to the distribution of scores in the general U.S. population. Specifically, all scores above and below 50 are above and below the average, respectively, in the 1998 general U.S. population. Because the standard deviation is 10 for both PCS and MCS measures, each one point difference in scores also has a direct interpretation. A one point difference is one-tenth of a standard deviation.

Table 7.2   1998 General U.S. Population Means, Standard Deviations and 1990 Factor Score Coefficients Used to Derive PCS and MCS Scale Scores, Acute Form

SF-36® Scale	Mean*	Standard Deviation*	Factor Score Coefficients	
			PCS	MCS
PF	82.62455	24.43176	0.42402	-0.22999
RP	82.65109	26.19282	0.35119	-0.12329
BP	73.86999	24.00884	0.31754	-0.09731
GH	70.78372	21.28902	0.24954	-0.01571
VT	58.41968	20.87823	0.02877	0.23534
SF	85.11568	23.24464	-0.00753	0.26876
RE	87.50009	22.01216	-0.19206	0.43407
MH	75.76034	18.04746	-0.22069	0.48581

*Note: The means and standard deviations for each SF-36v2™ are based on the 0-100 scoring.

Steps in Scoring

Following the scoring of the eight scales according to the standard SF-36v2scoring algorithms (0-100 scale) explained earlier in chapter 6, PCS and MCS are scored in three steps as explained below:

Standardization of Scales (Z-Scores), Acute Form

The first step in computing PCS and MCS consists of standardizing each of the 8 SF-36 scales using a z-score transformation. This is the same as Step 1 used in the norm-based scoring of the 8 SF-36 scales explained in Chapter 6. A z-score for each scale is computed by subtracting the mean 0-100 general US population score (see Table 7.2) for each SF-36 scale and dividing the difference by the corresponding scale standard deviation. Note that the SF-36 scales scored on the 0-100 scale are used in step 1. Norm-based SF-36 scale scores are not used in this step. Formulas are listed below.

Step 1. Formulas for z-score standardization of SF-36v2™ scales (Acute version)

$$PF\_Z = (PF -82.62455) / 24.43176$$
$$RP\_Z = (RP - 82.65109) / 26.19282$$
$$BP\_Z = (BP - 73.86999) / 24.00884$$
$$GH\_Z = (GH - 70.78372) / 21.28902$$
$$VT\_Z = (VT - 58.41968) / 20.87823$$
$$SF\_Z = (SF - 85.11568) / 23.24464$$
$$RE\_Z = (RE - 87.50009) / 22.01216$$
$$MH\_Z = (MH - 75.76034) / 18.04746$$

Means and standard deviations are from Table 7.2

Aggregation of Scale Scores (Acute Form)

After a z-score has been computed for each SF-36v2 scale, the second step involves computation of aggregate scores for the physical and mental components using the physical and mental factor score coefficients from the 1990 general U.S. population as given in Table 7.2.

Computation of an aggregate physical component score consists of multiplying each SF-36v2 scale z-score by its respective physical factor score coefficient and summing the eight products, as shown in step 2. Similarly, an aggregate mental component score is obtained by multiplying each SF-36 scale z-score by its respective mental factor score coefficient and summing the eight products.

Step 2. Formulas for aggregating standardized scales in estimating aggregate physical and mental component scores (Acute Form)

$AGG\_PHYS = (PF\_Z * .42402) + (RP\_Z * .35119) + (BP\_Z * .31754) + (GH\_Z * .24954) + (VT\_Z * .02877) + (SF\_Z * -.00753) + (RE\_Z * -.19206) + (MH\_Z * -.22069)$

$AGG\_MENT = (PF\_Z * -.22999) + (RP\_Z * -.12329) + (BP\_Z * -.09731) + (GH\_Z * -.01571) + (VT\_Z * .23534) + (SF\_Z * .26876) + (RE\_Z * .43407) + (MH\_Z * .48581)$

Transformation of Summary Scores

The third step involves transforming each component score to the norm-based (50, 10) scoring. This is accomplished by multiplying each aggregate component scale score by 10 and adding the resulting product to 50. Formulas are shown in step 3.

Step 3. Formulas for T-score transformation of component scores

Transformed Physical (PCS) = 50 + (AGG_PHYS * 10)

Transformed Mental (MCS) = 50 + (AGG_MENT * 10)

Missing Data Estimation

Results from ongoing evaluations of options to score PCS and MCS when a respondent is missing any one of the eight SF-36v2 scales has shown considerable promise. Evaluation of missing data rates across general and clinical populations has shown that 50% of those who have missing PCS and MCS scores are missing PCS and MCS because of missing data on one SF-36 scale. QualityMetric's Missing Data Estimator (MDE) will enable you to score PCS and MCS with data for 7 of the 8 SF-36 scales. PCS and MCS scores from MDE have proven to be reliable and valid (Ware et al., forthcoming). Further efforts are underway to evaluate options for scoring PCS and MCS when more than 1 SF-36 scale score is missing. More information will be available at these websites www.sf-36.com and www.qmetric.com.

Until QualityMetric's MDE is made available, it is recommended that component scale scores be set to missing if the respondent is missing any one of the eight SF-36v2 scales. To minimize the number of component scores missing, we recommend estimating each of the eight scale scores if half or more of the items are complete, as documented earlier.

Features of PCS and MCS Scores

The PCS and MCS were constructed and scored to achieve a number of advantages, in addition to reducing the SF-36v2™ from an eight-scale profile to two summary measures without substantial loss of information. Features of the PCS and MCS scores, including their reliability, confidence intervals (CI), skewness (percent ceiling and floor), and number of levels observed in the general U.S. population, are summarized in Table 7.3. These results confirm some of the theoretical advantages of the two summary measures as compared to the eight SF-36 scales, including a very large increase in the number of levels defined, smaller confidence intervals relative to each of the eight scales, as well as the elimination of both floor and ceiling effects.

Table 7.3   Comparison of Features of SF-36v2™ Scales and Summary Measures, 1998 General U.S. Population

	SF-36v2™ Scales[b]	Summary Measures	
		PCS	MCS
Reliability	.84 - .95	.95	.93
95% CI (±)	12 – 17	4.5	5.2
% Floor	0.0 – 2.1	0	0
% Ceiling	2.1 – 60.2	0	0
# of Levels[a]	8 - 20	486	494

[a] *Scores rounded to first decimal place*
[b] *Statistics are presented as the range of results found across the eight SF-36v2™ scales (standard and acute forms) in the 1998 general U.S. population.*

Scoring Checks

Because errors can lead to inaccurate scale scores, we strongly recommend formal scoring checks of SF-36v2 scales prior to computing the SF-36 component summary scales. These formal scoring checks are explained in full detail in chapter 6: scoring SF-36 scales.

The following scoring checks are also strongly recommended for the SF-36v2 component summary scales. Any discrepancies should be investigated for scoring errors:

Check correlation's between the eight SF-36v2 scales and the PCS and MCS scales. The PF, RP, and BP scales should correlate highest with the PCS and lowest with the MCS. The MH, RE, and SF scales should correlate highest with the MCS and lowest with the PCS. The GH and VT scales should correlate moderately with both physical and mental component scales.

Check the correlation between the physical and mental component summary scales. The correlation should be very low.

Scoring Exercise

The algorithms for scoring the SF-36v2 scales and summary measures for use with SAS software (along with a test data set) are available in electronic format to purchasers of this manual – by sending an email request to info@qmetric.com. QualityMetric also offers a Scoring & Data Quality Analysis service (information available at www.qmetric.com).

Table 7.4 presents descriptive statistics for the eight SF-36v2 scales (0-100 and norm-based) and the physical and mental component summary measures from the test dataset for the standard form. Table 7.5 presents descriptive statistics for the eight SF-36v2 scales (0-100 and norm-based) and the physical and mental component summary measures from the test dataset for the acute form.

After scoring the test dataset, you should observe the same means, standard deviations, and minimum and maximum observed values as those presented in Table 7.4 for the standard form and in Table 7.5 for the acute form.

Table 7.4   Test Dataset Descriptive Statistics˙: SF-36v2™ Scales and Summary Measures (N = 100), Standard Form

	Number of Cases	0-100 Scores				Norm-Based Scores			
		Mean	Standard Deviation	Min. Value	Max. Value	Mean	Standard Deviation	Min. Value	Max. Value
Physical Functioning	100	82.16	22.00	15	100	49.52	9.26	21.25	57.03
Role Physical	100	82.95	23.29	0	100	50.17	9.12	17.66	56.85
Bodily Pain	100	71.53	20.18	22	100	50.08	8.53	29.15	62.11
General Health	100	71.39	17.89	10	100	50.25	8.53	20.99	63.89
Vitality	100	60.35	18.13	7	94	51.01	9.06	23.99	67.70
Social Functioning	100	91.00	16.95	25	100	52.92	7.39	24.12	56.84
Role Emotional	100	89.50	18.93	0	100	50.98	8.83	9.23	55.87
Mental Health	100	78.86	14.85	40	100	52.18	8.36	30.29	64.08
Physical Component Summary (PCS)	100	N/a	-	-	-	49.29	8.66	23.93	61.58
Mental Component Summary (MCS)	100	N/a	-	-	-	52.58	7.71	30.06	64.18

˙*All descriptive statistics are rounded to the second decimal point.*

Table 7.5 Test Dataset Descriptive Statistics*: SF-36v2™ Scales and Summary Measures (N = 100), Acute Form

	Number of Cases	0-100 Scores				Norm-Based Scores			
		Mean	Standard Deviation	Min. Value	Max. Value	Mean	Standard Deviation	Min. Value	Max. Value
Physical Functioning	100	82.44	24.19	0	100	49.92	9.90	16.18	57.11
Role Physical	100	82.91	26.66	0	100	50.10	10.18	18.44	56.62
Bodily Pain	100	71.8	24.56	12	100	49.13	10.23	24.23	60.88
General Health	100	70.97	19.89	5	100	50.08	9.34	19.09	63.72
Vitality	100	58.35	19.93	7	100	49.96	9.54	25.01	69.91
Social Functioning	100	85.87	23.68	0	100	50.32	10.18	13.38	56.40
Role Emotional	100	90.16	19.76	0	100	51.21	8.97	10.24	55.67
Mental Health	100	77.55	15.52	20	100	50.99	8.60	19.10	63.43
Physical Component Summary (PCS)	100	N/a	-	-	-	49.29	9.33	19.34	65.35
Mental Component Summary (MCS)	100	N/a	-	-	-	51.17	8.73	16.38	

All descriptive statistics are rounded to the second decimal point.

SF-36v2 algorithms have been made available to computer software vendors and other organizations providing scoring and analysis services for the SF-36. Look for the symbol to the right.

This symbol is your assurance that computer software products and data processing services produce results that are comparable with this Manual and with other normative data and interpretation guidelines for the SF-36v2 Health Survey.

SF-36v2™