Interpreting the SF-12

Interpreting the SF-12

The results for this report are based on the responses of 7,520 respondents in the 2001 Utah Health Status Survey (HSS). The SF-12 is a multipurpose short form survey with 12 questions, all selected from the SF-36 Health Survey (Ware, Kosinski, and Keller, 1996). The questions were combined, scored, and weighted to create two scales that provide glimpses into mental and physical functioning and overall health-related-quality of life.

The SF-12 is a generic measure and does not target a specific age or disease group. It has been developed to provide a shorter, yet valid alternative to the SF-36, which has been seen by many health researchers as too long to administer to studies with large samples. The SF-12 is weighted and summed to provide easily interpretable scales for physical and mental health.

Physical and Mental Health Composite Scores (PCS & MCS) are computed using the scores of twelve questions and range from 0 to 100, where a zero score indicates the lowest level of health measured by the scales and 100 indicates the highest level of health. The data obtained with the SF-12 has been developed, tested and validated by Quality Metric Incorporated. The 2001 HSS uses the second version of the SF-12, rather than the first version, which was used for the 1996 Health Status Survey. There are minor differences in terms of wording and scoring between the two versions; however, they are summed and weighted to be comparable with each other.

The Physical and Mental Health Composite Scale scores (PCS & MCS) derived from the SF-12 have little intuitive meaning. This is because the range of possible scores varies considerably. PCS and MCS scores tend to vary over the life span for different age groups as well [PCS tends to decrease with age, while MCS tends to increase]. It would not make sense to say that a PCS score of 45.3 means the same thing for a person who is 25 years old compared to a person who is 65 years old. It is because of this inherent variation SF-12 scores have over the lifespan that it is useful to introduce the idea of age-specific mean difference scores.

The age-specific mean difference score (difference score) is the amount by which a person's score differs from their age group's mean score. In other words, an individual with a difference score of -5.5 has scored 5.5 points lower than the mean score for their age group, indicating somewhat poorer health. By looking at difference scores, it is clear whether a person is more or less healthy than other persons in his or her comparison group. For individual scores, those that score higher than the mean indicate a person has better health status than most others their age. Conversely, scores that are lower than the mean indicate a person has poorer health than most others their age. A key advantage of age-specific mean difference scores is that a difference score of -5.5 means the same thing in terms of relative health for a person regardless of age. Another advantage of using difference scores is the ability to compare the association that different sociodemographic and disease/lifestyle factors have with physical and mental health. This report will use age-specific difference scores extensively because they are easier to interpret than the PCS and MCS scores.

Interpreting the SF-12: Comparing Versions 1 and 2 of the SF-12

The Utah Health Status Survey 2001 used an updated version (Version 2) of the SF-12 to measure health status. Version 2 of the SF-12 differs from Version 1 in several ways. Changes in the administration of the SF-12 are based on more than 10 years of experience with findings reported in thousands of publications based on the SF-36 and SF-12 (Version 1) Health Surveys (Ware et al., 2002). A brief description of similarities and differences between the two versions is included below:

- Both versions use the same basic 12 questions to measure physical and mental health status. However, changes were made to question wording, instructions and formatting for the second version;
- These changes in the layout and response category options for the second version are meant to make it easier to read and complete the questions, thereby reducing missing responses;
- The second version is designed to provide greater comparability with translations and cultural adaptations that are widely-used in the U.S. and other countries;
- Four items in the second version were changed from dichotomous to five-level response categories;
- Six-level response categories were changed to five-level response categories to simplify items in the Mental Health and Vitality scales.

Frequency distribution bar charts from the 2001 HSS for the second version of the SF-12 are provided below. Question wording and response categories for the first version of the SF-12 are also provided for comparison.

General Health Subdomain

Version 1: In general, would you say your health is excellent, very good, good, fair, or poor?
1) Excellent
2) Very good
3) Good
4) Fair
5) Poor

Version 2: In general, would you say your health is excellent, very good, good, fair or poor?

Percentage of Persons

| Poor	Fair	Good	Very Good	Excellent
2.9%	9.5%	24.8%	33.8%	29.0%
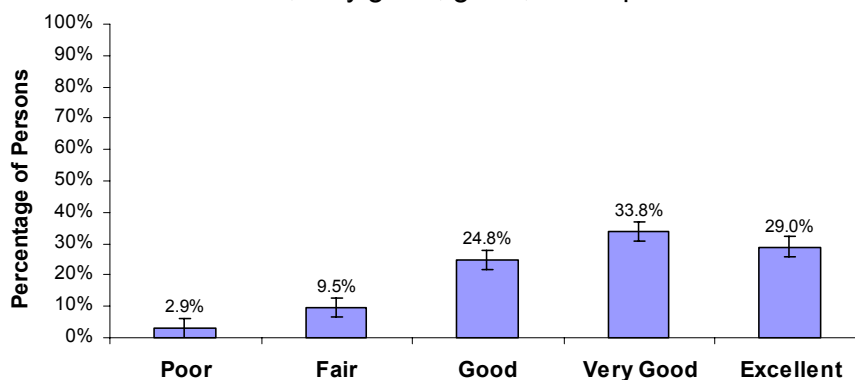
Physical Functioning Subdomain

Version 1: How does your health now limit you in moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf? Would you say you are limited a lot, a little or not at all?

    1) Yes, limited a lot
    2) Yes, limited a little,
   3) No, not limited at all.

Version 2: Are you now limited in moderate activities, such as moving a table, pushing a vacuum cleaner, bowling or playing golf? Does your health now limit you a lot, limit you a little or not limit you at all?

Response	Percentage of Persons
A lot	7.6%
A little	12.5%
Not at all	79.8%
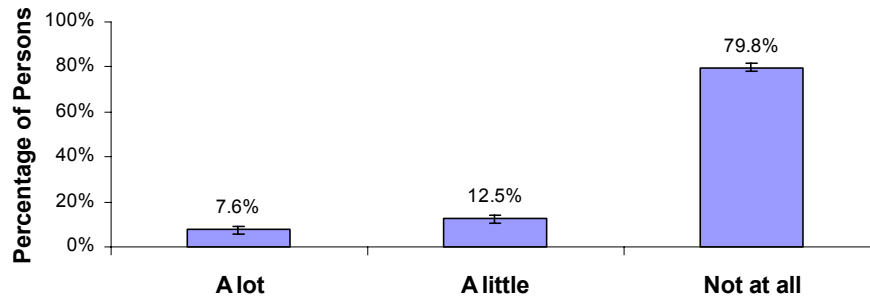
Version 1: How about climbing several flights of stairs? Would you say your health limits you a lot, a little, or not at all?

    1) Yes, limited a lot
    2) Yes, limited a little,
    3) No, not limited at all.

Version 2: How about climbing several flights of stairs? Would you say your health now limits you a lot, limits you a little, or does not limit you at all?

| Response	Percentage of Persons
A lot	8.3%
A little	16.2%
Not at all	75.5%
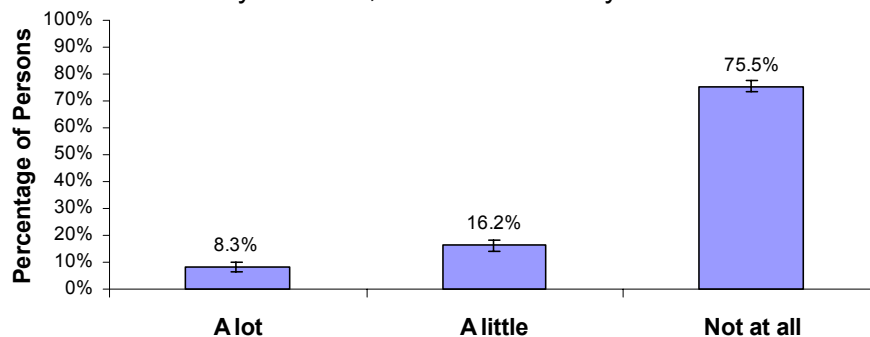
Role Functioning (Physical) Subdomain

Version 1: Thinking about the past four weeks, have you accomplished less than you would like as a result of your physical health?

    1) Yes
    2) No

Version 2: During the past 4 weeks, how much of the time have you had any of the following problems with your work or regular daily activities as a result of your physical health? How much of the time have you accomplished less than you would like?
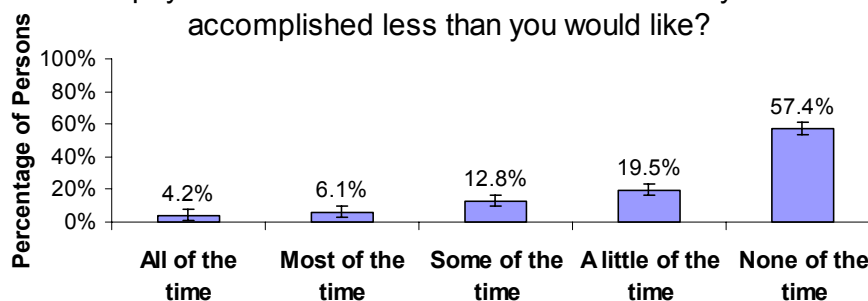
Version 1: During the past four weeks, were you limited in the kind of work or other activities you could do as a result of your physical health?

    1) Yes
    2) No

Version 2: How much of the time were you limited in the kind of work or other activities you could do?
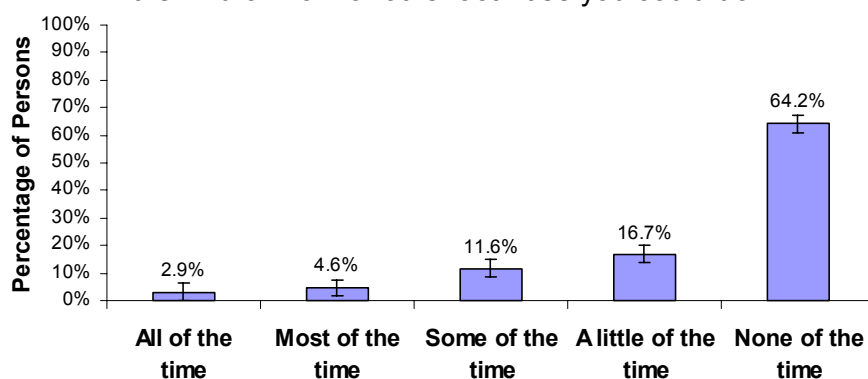
Bodily Pain Subdomain

Version 1: During the past four weeks, how much did pain interfere with your normal work including both work outside the home and housework?
> 1) Extremely
> 2) Quite a bit
> 3) Moderately
> 4) A little bit
> 5) Not at all

Version 2: During the past four weeks, how much did pain interfere with your normal work including both outside the home and housework, would you say...?
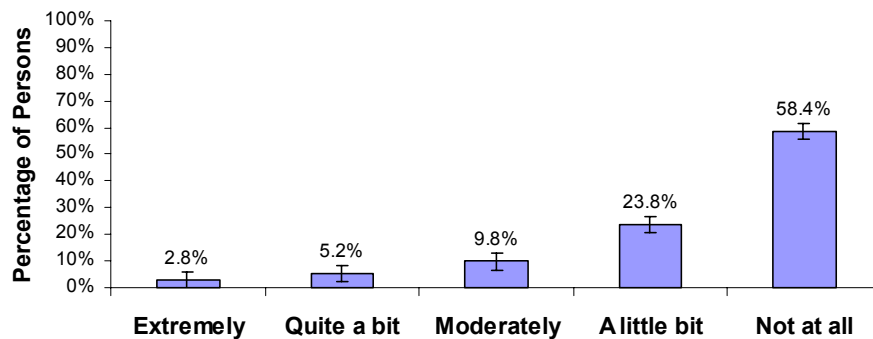

Vitality Subdomain

Version 1: How much of the time during the past four weeks did you have a lot of energy? Would you say (read responses)?
> 1) None of the time
> 2) A little of the time
> 3) Some of the time
> 4) Good bit of the time
> 5) Most of the time
> 6) All of the time

Version 2: How much of the time during the past four weeks did you have a lot of energy? Would you say...?

Percentage of Persons

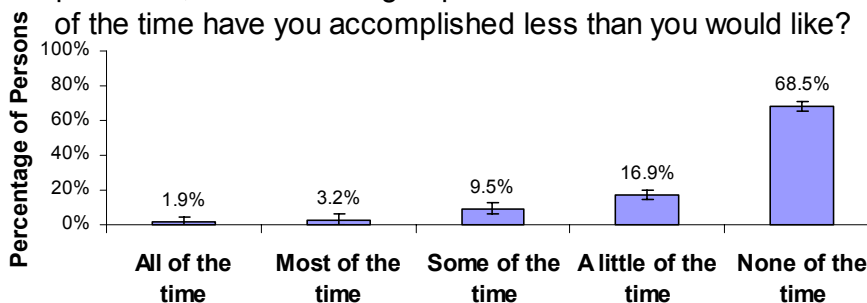|	All of the time	Most of the time	Some of the time	A little of the time	None of the time
	3.2%	9.0%	25.0%	50.7%	12.2%
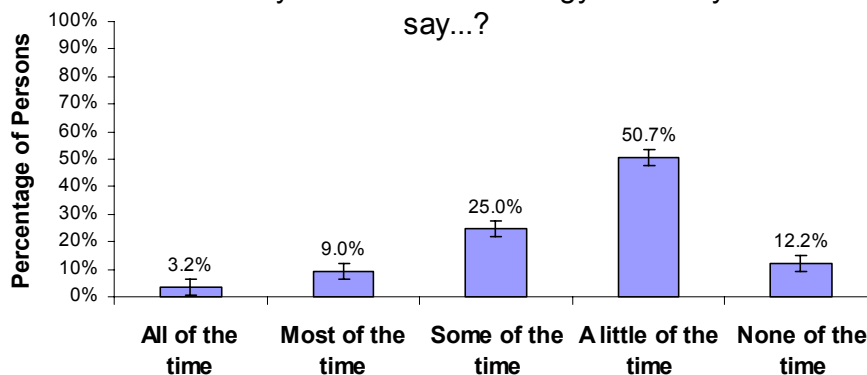
Role Functioning (Emotional) Subdomain

Version 1: In the past four weeks, did you accomplish less than you would like as a result of an emotional problem, such as feeling depressed or anxious?
  1) Yes
  2) No

Version 2: During the past four weeks, how much of the time have you had any of the following problems with your work or other daily activities as a result of any emotional problems, such as feeling depressed or anxious. How much of the time have you accomplished less than you would like?

Percentage of Persons

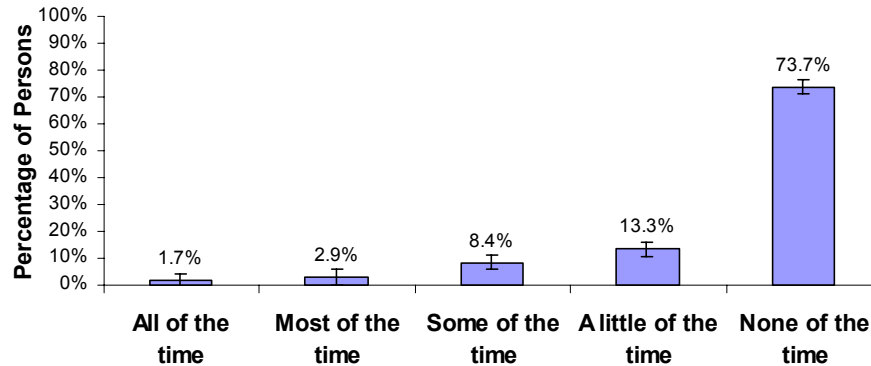|	All of the time	Most of the time	Some of the time	A little of the time	None of the time
	1.9%	3.2%	9.5%	16.9%	68.5%

Version 1: During the last four weeks, did you have trouble doing work or other activities as carefully as usual as a result of an emotional problem, such as feeling depressed or anxious?
  1) Yes
  2) No

Version 2: How much of the time did you have trouble doing work or other activities as carefully as usual?

Percentage of Persons

- All of the time: 1.7%
- Most of the time: 2.9%
- Some of the time: 8.4%
- A little of the time: 13.3%
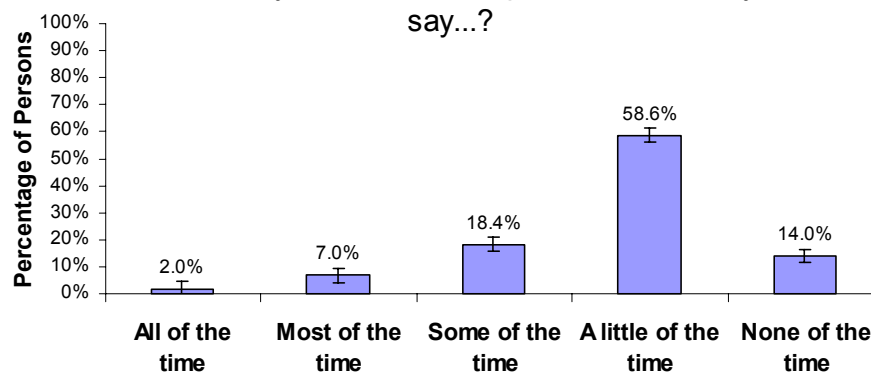- None of the time: 73.7%

Mental Health Subdomain

Version 1: How much of the time during the past four weeks have you felt calm and peaceful? Would you say (read responses)?
1) None of the time
2) A little of the time
3) Some of the time
4) Good bit of the time
5) Most of the time
6) All of the time

Version 2: How much of the time during the past four weeks have you felt calm and peaceful? Would you say...?

Percentage of Persons

- All of the time: 2.0%
- Most of the time: 7.0%
- Some of the time: 18.4%
- A little of the time: 58.6%
- None of the time: 14.0%

Version 1: How much of the time during the past four weeks have you felt downhearted and blue? (If necessary, read responses)

> 1) All of the time
> 2) Most of the time
> 3) Good bit of the time
> 4) Some of the time
> 5) A little of the time
> 6) None of the time

Version 2: How much of the time during the past four weeks have you felt downhearted and blue?
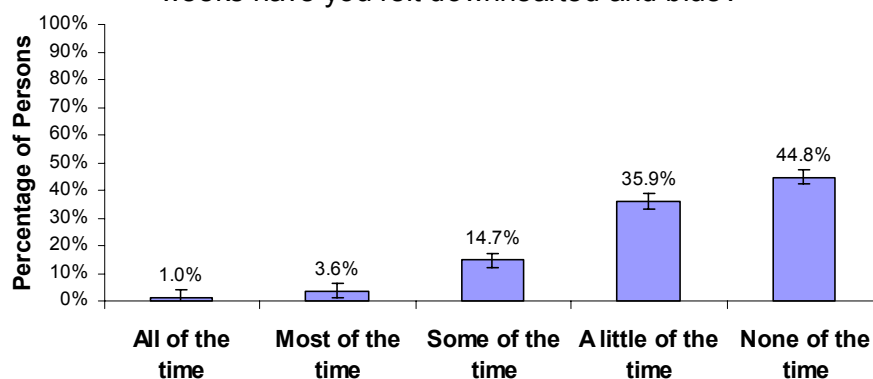
Social Functioning Subdomain

Version 1: During the last four weeks, how much of the time has your physical health or emotional problems interfered with your social activities, like visiting with friends, relatives etc.? (If necessary, read responses)

Version 2: During the last four weeks, how much of the time has your physical health or emotional problems interfered with your social activities, like visiting with friends, relatives, etc.?
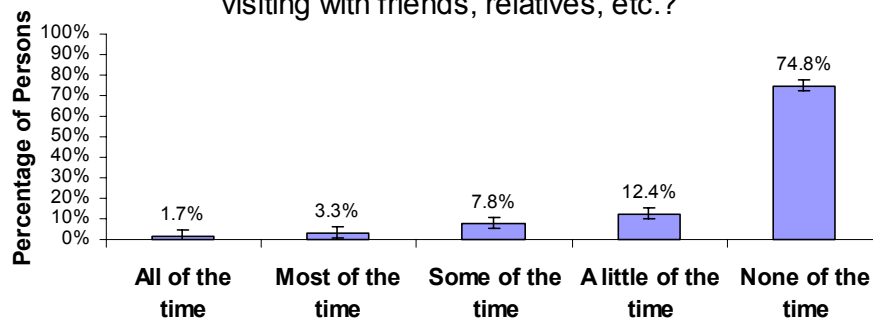
Interpreting the SF-12: Composite Scales

Medical Outcomes Study SF-12 Physical Health Composite Scale Scores: Utahns Age 18 or Over, 2001

Mean
50.8

Percentage of Persons With Score (y-axis)

SF-12 Physical Health Composite Scale Score (x-axis)

Medical Outcomes Study SF-12 Mental Health Composite Scale Scores: Utahns Age 18 or Over, 2001

Mean
52.4

Percentage of Persons With Score (y-axis)
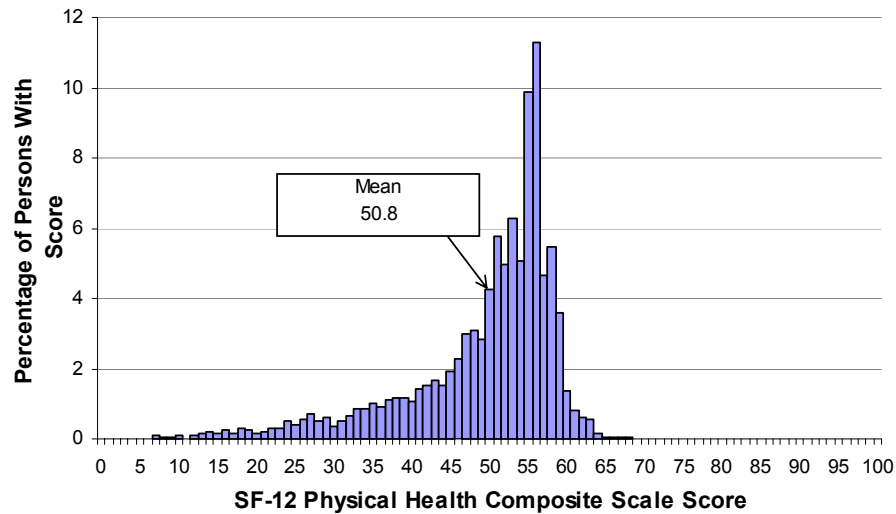
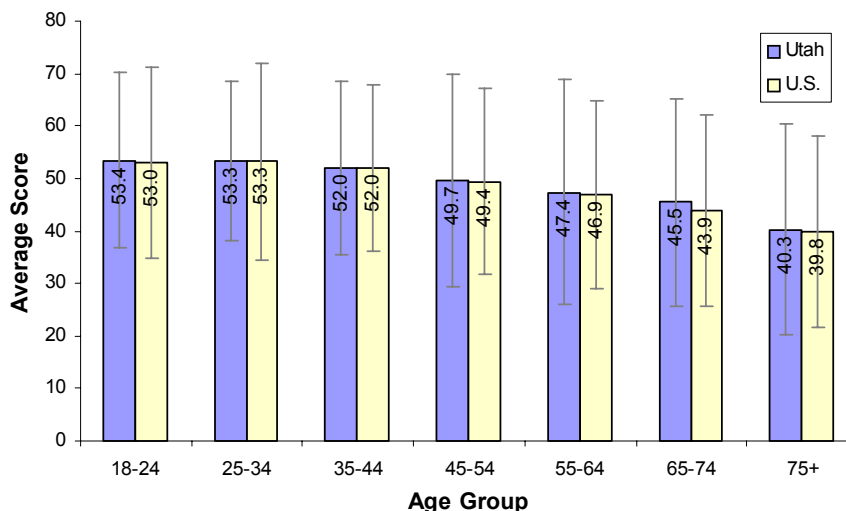**SF-12 Mental Health Composite Scale Score** (x-axis)
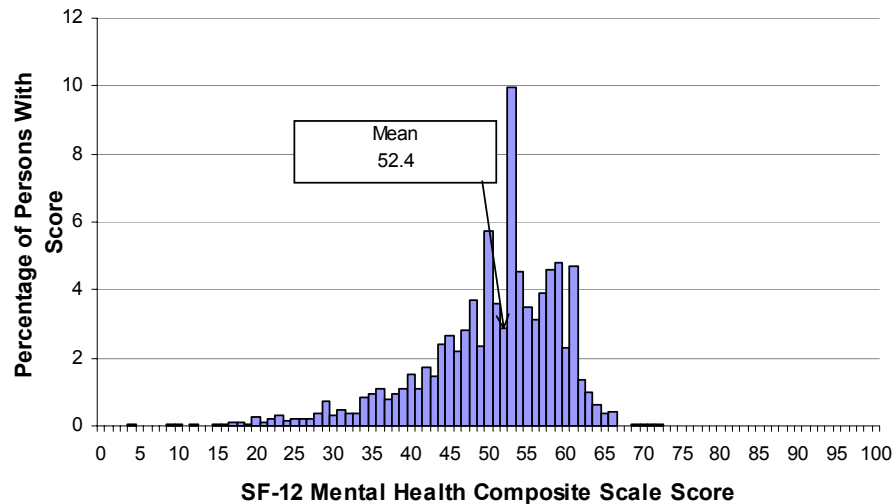
- Physical and Mental Health Composite Scores (PCS & MCS) are computed using the scores of twelve questions and range from 0 to 100, where a zero score indicates the lowest level of health measured by the scales and 100 indicates the highest level of health. The histograms above illustrate the distribution of composite scale scores.

- Both Physical and Mental Health Composite Scales combine the 12 items in such a way that they compare to a national norm with a mean score of 50.0 and a standard deviation of 10.0.

- In Utah, the mean (average) scores are 50.8 for the PCS, and 52.4 for the MCS.
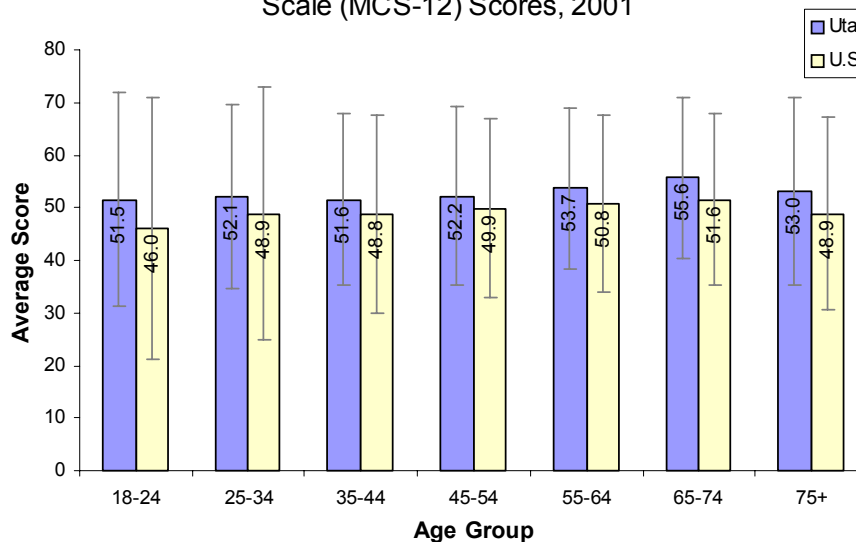
Comparison of Utah and U.S. Physical Health Composite Scale (PCS-12) Scores, 2001

Comparison of Utah and U.S. Mental Health Composite Scale (MCS-12) Scores, 2001
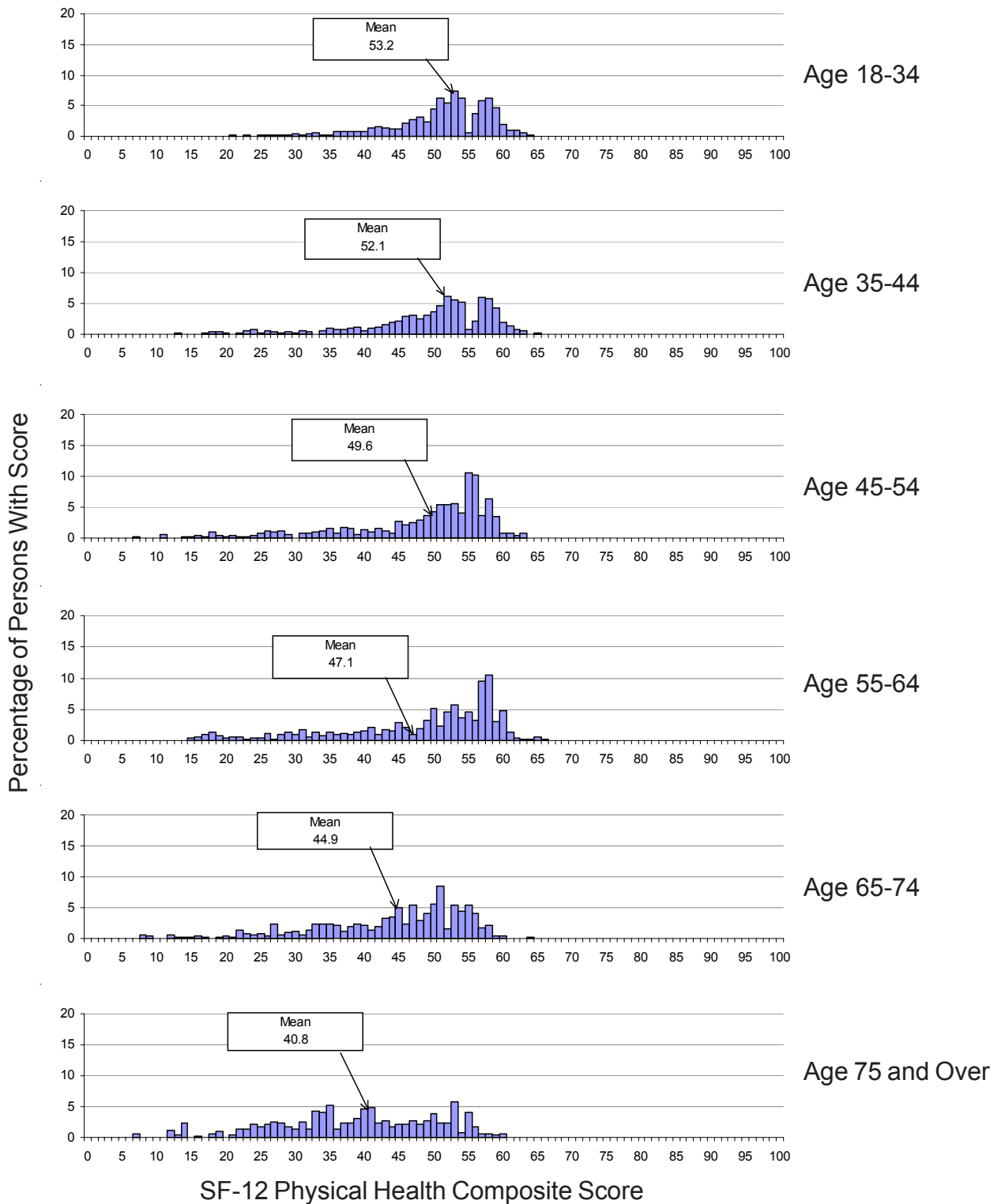

- With age, persons tend to score lower on the physical health scale (PCS) but higher on the mental health scale (MCS). Because there are systematic age differences, it is important to interpret a person's score in the context of other persons near their age.

- Utahns of all age groups scored higher than the U.S. norm on the mental health scale (MCS).

Medical Outcomes Study SF-12 Physical Health Composite Scale Scores by Age Group, Utahns Age 18 or Over, 2001

Percentage of Persons With Score

SF-12 Physical Health Composite Score

Medical Outcomes Study SF-12 Mental Health Composite Scale Scores by Age Group, Utahns Age 18 or Over, 2001

Age 18-34

Mean 51.8

Age 35-44

Mean 51.2

Age 45-54

Mean 52.2

Age 55-64

Mean 53.9

Age 65-74

Mean 54.8

Age 75 and Over

Mean 52.6

Percentage of Persons With Score

SF-12 Mental Health Composite Score

Computing Difference Scores

- Difference scores can be used to help interpret the meaning of scale values. The difference score is the difference between a person's score and the mean or average score for his or her age group.
- A positive score means the person is healthier than average. A negative score means a person is less healthy than average.

Age 18-34

PCS-12 Score

PCS Difference Score

Below Average   Average   Above Average
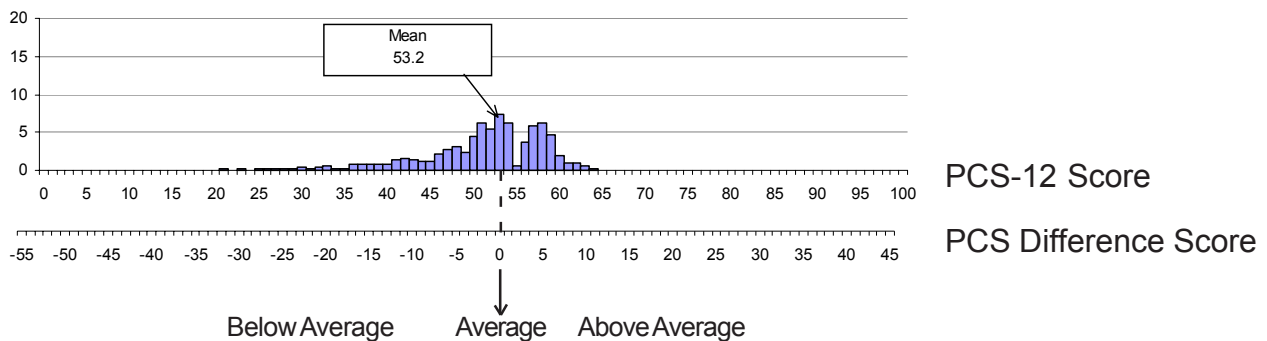
Establishing Cut-off Points for Exceptionally Good and Poor Health

In addition to knowing whether a person's score is above or below average, it is also helpful to know whether a person's score is significantly above or below average. If a person's physical health difference score is negative but close to zero, he or she should probably be considered in 'average' health. However, if a person's physical health difference score is hovering around negative 20, he or she should probably be considered 'below average,' or in poor health. This section will discuss the methods used to derive cut-off points for defining where average health ends and below or above average health begins.

Statistical Methods for Establishing Cut-off Points

This report used simple statistical methods to establish cut-off points for average and below average health. These methods rely on measures of variability (such as standard deviation and standard error) and use confidence intervals to define the average level of health. If the confidence interval for a person or group's score includes the zero point (average score), then the score does not differ from the average. However, if the confidence interval does not include the zero point, then the score is different from the average (either above or below average).

Cut-off Points for Individual Scores. The cut-off for an individual's score is based on a property of the SF-12 scale called the standard error of measurement. Over multiple administrations of the survey, some degree of variation in a person's PCS or MCS would be expected and would not necessarily reflect a change in the individual's health status. The 95% confidence interval is calculated as 1.96 times the standard error of measurement. The calculated 95% confidence interval for the Physical Composite Scale (and PCS Difference Score) is $\pm$ 6.97; and for the Mental Composite Scale (and MCS Difference Score) is $\pm$ 6.24. In other words, an

individual's PCS score plus or minus 6.97 gives the range in which the person's score is likely to fall 95% of the time, providing no major changes in health status occur. If this range of values (for the PCS) includes the zero point, then that person's health is considered to be average for his or her age group.

To demonstrate further, if we apply this approach to an individual's PCS difference score, then the 95% confidence interval for a score that is below –6.97 will not include zero and will be considered 'below average' by this criterion. In order to further illustrate, it is useful to consider several hypothetical cases (see Example 1). Case 1 is a female with no chronic medical problems and a PCS difference score of +1.6, Case 2 is a male has been diagnosed with diabetes and has a PCS difference score of –4.51, Case 3 is a female with chronic obstructive pulmonary disease and a PCS difference score of –16.2, and Case 4 is a male with asthma and a PCS difference score of –7.23. Cases 1 and 2 are not significantly different from the average score (the confidence intervals around their scores include the zero point). Cases 3 and 4, however, are significantly below the average (they do not include the zero point).

Example 1. Physical Health Status Difference Scores and Confidence Intervals for Four Hypothetical Individuals
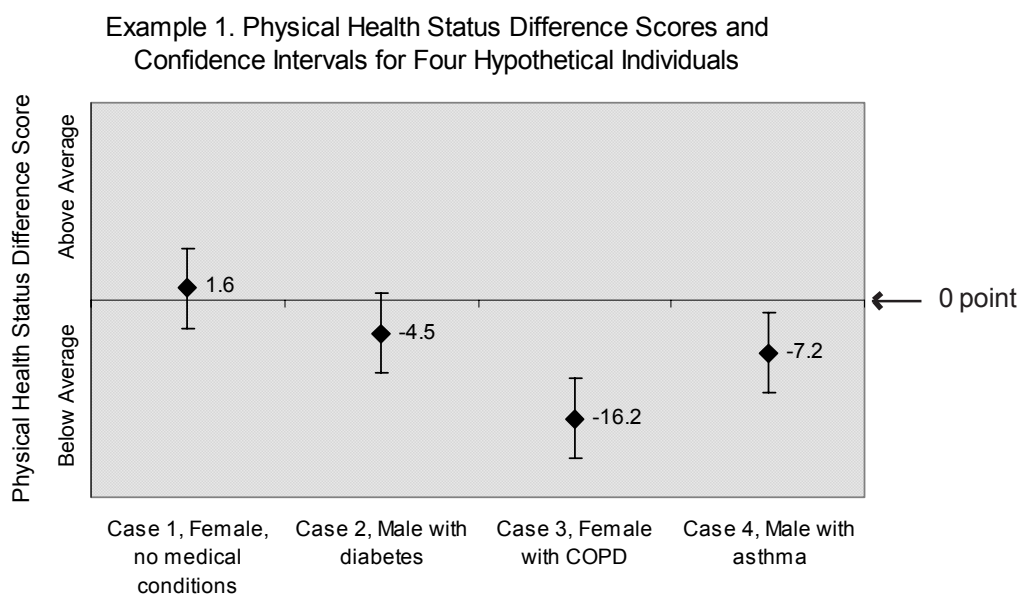
Cut-off Points for Group Means. The mean or average score has a measure of deviation, the standard error that is based on the amount of dispersion or spread the group's scores around the mean score and the number of persons in that group. Example 2 (below) shows group means and standard errors that have been plotted for males and females by income category. Both males and females living in household with less than $20,000 annual income scored significantly below the average (the confidence interval does not include zero), while those in the middle two income ranges scored in the average range, and males and females with $65,000 or more annual income had PCS difference scores that were significantly above the average.

Example 2. Physical Health Status Difference Scores and
Confidence Intervals for Income by Sex Means

Legend:
- Under $20,000
- $20,000 - <$45,000
- $45,000 - <$65,000
- $65,000 and Over

Physical Health Status Difference Score (Above Average / Below Average)

Male:
- -3.1
- -0.6
- 0.9
- 2.2

Female:
- -4.9
- -1.3
- -0.8
- 1.8