



Dept. of Population Health Sciences  
University of Wisconsin-Madison

# **A Bayesian Estimation of an Average SF-6D Preference-Based Score from Commonly Reported SF-12 Statistics**

Janel Hanmer, PhD Candidate, Dennis G Fryback, PhD  
University of Wisconsin – Madison  
ISPOR 11<sup>th</sup> Annual International Meeting  
May 22<sup>nd</sup>, 2006

[jehanmer@wisc.edu](mailto:jehanmer@wisc.edu), [dfryback@wisc.edu](mailto:dfryback@wisc.edu)

This work was funded by a NIH Integrated Training for Physician Scientists Training Grant (GM008692), an AHRQ T-32 Pre-doctoral NRSA Training Grant (HS00083), and P01AG20679 from the National Institute on Aging (“Norms and Performance Comparisons for 5 Health Indexes”)



Dept. of Population Health Sciences  
University of Wisconsin-Madison

# The SF-12 Health Measurement System<sup>1</sup>

- ❑ 12 multiple choice questions
- ❑ A standardized subset of the SF-36
- ❑ Chosen to most closely estimate the Mental Component Score (MCS) and Physical Component Score (PCS) from the SF-36
  - The MCS and PCS have means of 50 with a standard deviation of 10
- ❑ Widely used (610 citations in Medline)

1. Ware JE, Kosinski M, Keller SD. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scores. Boston, MA: The Health Institute, New England Medical Center, March 1995



Dept. of Population Health Sciences  
University of Wisconsin-Madison

## **MCS and PCS Scores Cannot be Used in Cost-Effectiveness Analyses**

- ❑ US Panel on Cost-Effectiveness in Health and Medicine (1996) recommended a single, preference-based summary score of health for cost-effectiveness analyses
  
- ❑ There are algorithms which map MCS and PCS scores from the SF-12 to other preference-based instruments' summary scores:
  - Lawrence et al (2004), EQ-5D, MEPS
  - Franks et al (2004), EQ-5D, MEPS
  - Franks et al (2003), EQ-5D and HUI3, health center sample
  - Sengupta et al (2004), HUI3, managed care sample



Dept. of Population Health Sciences  
University of Wisconsin-Madison

# The SF-6D Scoring Algorithm

- ❑ Brazier et al<sup>2</sup> developed an algorithm which estimates a single, preference-based summary score using 7 questions from the SF-12
- ❑ Appropriate for use in cost-effectiveness analyses
- ❑ SF-6D values are between .345 and 1.0
- ❑ The SF-6D was developed 9 years after the development of the SF-12



Dept. of Population Health Sciences  
University of Wisconsin-Madison



*Quality of Life Research* 13: 449–456, 2004.  
© 2004 Kluwer Academic Publishers. Printed in the Netherlands.

449

## **Performance of the RAND-12 and SF-12 summary scores in type 2 diabetes**

Jeffrey A. Johnson<sup>1</sup> & Sheri L. Maddigan<sup>2</sup>

<sup>1</sup>*Department of Public Health Sciences, University of Alberta, Edmonton, Canada (E-mail: jeff.johnson@ualberta.ca);* <sup>2</sup>*Institute of Health Economics, Edmonton, Canada*



Dept. of Population Health Sciences  
University of Wisconsin-Madison

**Table 1.** Demographic and clinical characteristics (n = 372)

Age – Mean (SD)	62.27 (12.53)
Sex (males) – n (%)	164 (41.6)

**Table 2.** Comparison of mean (95% CI) mental and physical dimensions of HRQL for the SF-12 and RAND-12<sup>a</sup>

	Mental health		Physical health	
	MCS-12	RAND-12 MHC	PCS-12	RAND-12 PHC
Overall mean scores	48.49 (47.30–49.67) (n = 259)	44.92 (43.76–46.07) (n = 267)	41.98 (40.71–43.25) (n = 259)	43.52 (42.30–44.73) (n = 267)
Treatment regimen for diabetes				
Group 1 – diet only	50.57 (47.81–53.34)	46.73 <sup>b</sup> (44.05–49.42)	42.69 (39.70–45.66)	45.21 <sup>b</sup> (42.40–48.02)
Group 2 – oral medication	48.36 (46.70–50.02)	44.65 (42.76–46.55)	42.62 (40.82–44.41)	43.98 (42.07–46.03)
Group 3 – insulin ± oral medication	47.90 (45.18–50.61)	42.97 (40.53–45.40)	40.14 (37.20–43.07)	41.01 (38.55–43.67)
Duration of diabetes				
Less than or equal to 5 years	48.64 (47.59–50.02)	45.65 (43.79–47.52)	43.06 (40.90–45.22)	45.02 <sup>c</sup> (43.06–46.98)
Greater than 5 years	48.65 (39.30–50.96)	44.11 (42.12–46.10)	40.70 (38.02–43.39)	42.09 (39.97–44.22)



Dept. of Population Health Sciences  
University of Wisconsin-Madison

## Our Objective

To construct an algorithm which predicts the SF-6D preference-based score using commonly reported SF-12 variables and demographics.

Such an algorithm could allow comparisons to previously published results without requiring access to individual level data.



Dept. of Population Health Sciences  
University of Wisconsin-Madison

## Data Source: Medical Expenditures Panel Survey (MEPS)<sup>3</sup>

- ❑ MEPS is a set of surveys of non-institutionalized families and individuals in the United States
- ❑ Households containing Hispanics and blacks are oversampled
- ❑ SF-12 was mailed to all subjects aged 18 and older
- ❑ Model development :
  - 2002, SF-12 version 1, n=20,206
- ❑ Model evaluation:
  - 2001, SF-12 version 1, n=20,208
  - 2003, SF-12 version 2, n=20,368



Dept. of Population Health Sciences  
University of Wisconsin-Madison

# Model Selection

- ❑ SF-6D =  $f(\text{Age, Sex, MCS, PCS})$ 
  - linear models without interaction or power terms
  
- ❑ Fit using WinBUGS 1.4.1, a Bayesian modeling software package
  - The priors for all coefficients were diffuse normal
  
- ❑ Best model chosen using Deviance Information Criterion (DIC)



Dept. of Population Health Sciences  
University of Wisconsin-Madison

## The Best Fitting Model

$$\text{SF-6D} = -.00154 - .0022 * \text{Female} + .00014 * \text{Age} \\ + .0081 * \text{MCS} + .0081 * \text{PCS}$$

Where:

Female is 1 if the respondent is female, 0 if male

Age is the respondent's age in years

MCS is the respondent's MCS score

PCS is the respondent's PCS score



Dept. of Population Health Sciences  
University of Wisconsin-Madison

# Comparison to Other Algorithms

- ❑ Our model which predicts SF-6D summary scores:
  - $R^2 = .88$
  
- ❑ Models which predict EQ-5D summary scores:
  - $R^2 = .59$  Franks et al (03)
  - $R^2 = .63$  Franks et al (04)
  - $R^2 = .61$  Lawrence et al (04)
  
- ❑ Models which predict HUI3 summary scores:
  - $R^2 = .51$  Franks et al (03)
  - $R^2 = .55$  Sengupta et al (04)



Dept. of Population Health Sciences  
University of Wisconsin-Madison

## Average SF-6D Scores

To calculate an average SF-6D score:

$$\frac{1}{n} \sum_{i=1}^n \text{SF-6D}_i = \frac{1}{n} \sum_{i=1}^n (-.00154 - .0022 * \text{female}_i + .00014 * \text{age}_i + .0081 * \text{MCS}_i + .0081 * \text{PCS}_i)$$



## Average SF-6D Scores

To calculate an average SF-6D score:

$$\frac{1}{n} \sum_{i=1}^n \text{SF-6D}_i = \frac{1}{n} \sum_{i=1}^n (-.00154 - .0022 * \text{female}_i + .00014 * \text{age}_i + .0081 * \text{MCS}_i + .0081 * \text{PCS}_i)$$

A simple algebraic rearrangement removes the need for individual level data:

$$\frac{1}{n} \sum_{i=1}^n \text{SF-6D}_i = -.00154 - .0022 * \frac{1}{n} \sum_{i=1}^n (\text{female}_i) + .00014 * \frac{1}{n} \sum_{i=1}^n (\text{age}_i) + .0081 * \frac{1}{n} \sum_{i=1}^n (\text{MCS}_i) + .0081 * \frac{1}{n} \sum_{i=1}^n (\text{PCS}_i)$$



**Table 1.** Demographic and clinical characteristics (n = 372)

Age – Mean (SD)	62.27 (12.53)
Sex (males) – n (%)	164 (41.6)

**Table 2.** Comparison of mean (95% CI) mental and physical dimensions of HRQL for the SF-12 and RAND-12<sup>a</sup>

	Mental health		Physical health	
	MCS-12	RAND-12 MHC	PCS-12	RAND-12 PHC
Overall mean scores	48.49 (47.30–49.67) (n = 259)	44.92 (43.76–46.07) (n = 267)	41.98 (40.71–43.25) (n = 259)	43.52 (42.30–44.73) (n = 267)
<del>Treatment regimen for diabetes</del>				
Group 1 – diet only	50.57 (47.81–53.34)	46.73 <sup>b</sup> (44.05–49.42)	42.69 (39.70–45.66)	45.21 <sup>b</sup> (42.40–48.02)
Group 2 – oral medication	48.36 (46.70–50.02)	44.65 (42.76–46.55)	42.62 (40.82–44.41)	43.98 (42.07–46.03)
Group 3 – insulin ± oral medication	47.90 (45.18–50.61)	42.97 (40.53–45.40)	40.14 (37.20–43.07)	41.01 (38.55–43.67)
Duration of diabetes				
Less than or equal to 5 years	48.64 (47.59–50.02)	45.65 (43.79–47.52)	43.06 (40.90–45.22)	45.02 <sup>c</sup> (43.06–46.98)
Greater than 5 years	48.65 (39.30–50.96)	44.11 (42.12–46.10)	40.70 (38.02–43.39)	42.09 (39.97–44.22)



Dept. of Population Health Sciences  
University of Wisconsin-Madison

## Average SF-6D Estimate for the Example

$$\frac{1}{n} \sum_{i=1}^n \text{SF-6D}_i = -.00154 - .0022* .58 + .00014*62.27 \\ + .0081*50.57 + .0081*42.69$$

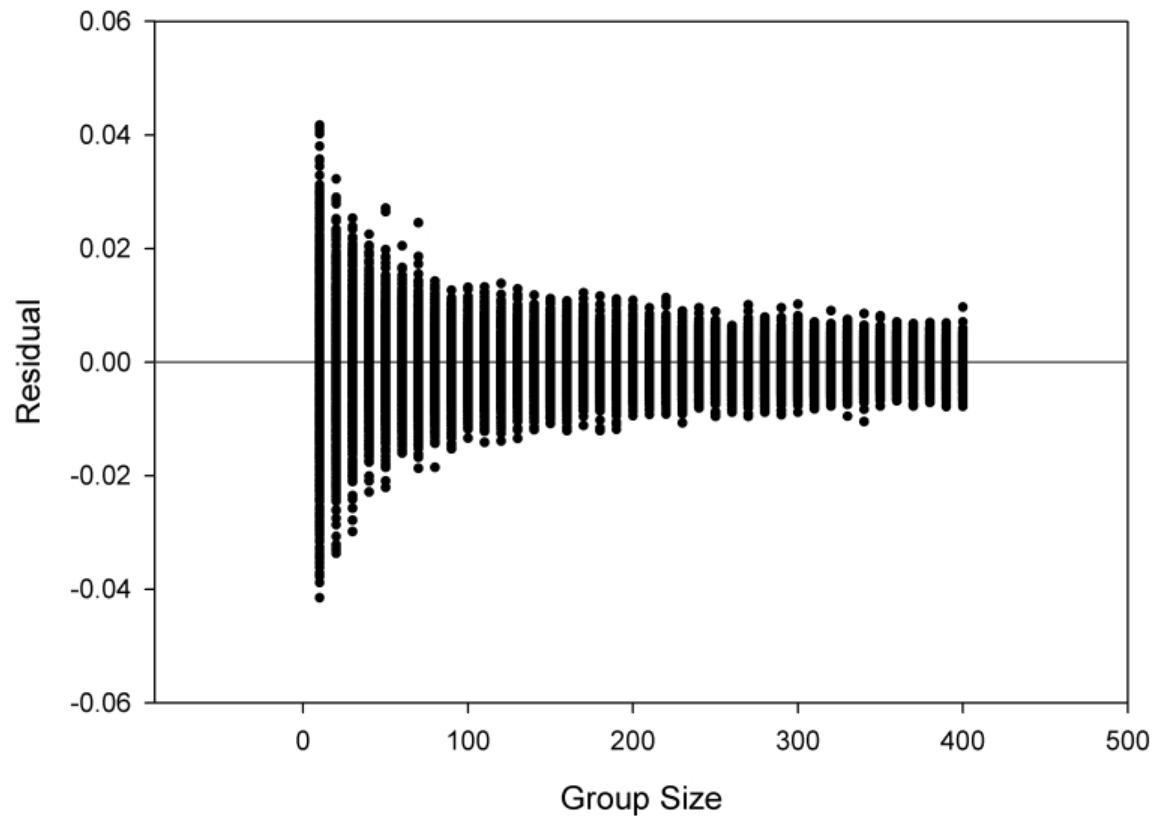
$$\frac{1}{n} \sum_{i=1}^n \text{SF-6D}_i = .761$$



Dept. of Population Health Sciences  
University of Wisconsin-Madison

# Error of Estimates Including Parameter Uncertainty

200 Residuals for Each Group Size



Standard Deviation of  
Residual:

0.014 for 10 subjects

0.006 for 50 subjects

0.005 for 100 subjects

0.003 for >200 subjects

Suggested 95%

Confidence Intervals:

$\pm 0.03$  for 10 subjects

$\pm 0.02$  for 50 subjects

$\pm 0.01$  for 100 subjects

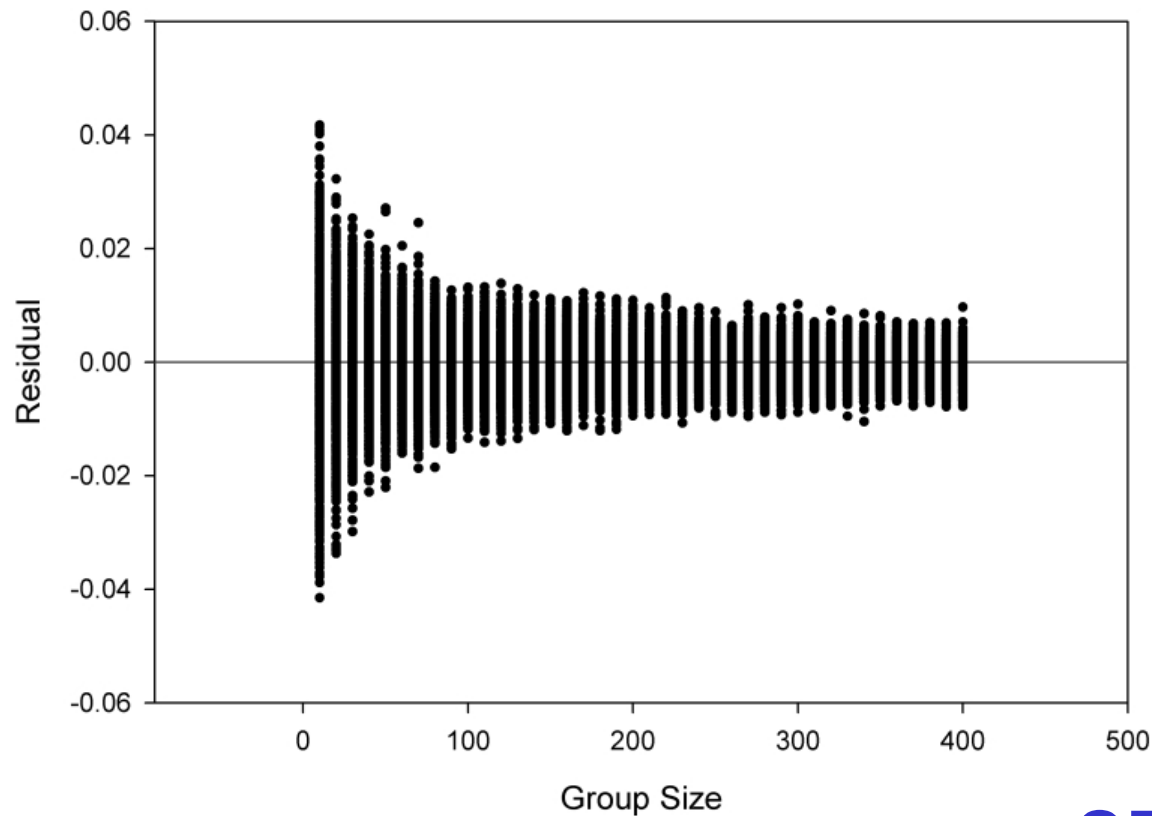
$\pm 0.006$  for >200 subjects



Dept. of Population Health Sciences  
University of Wisconsin-Madison

# Error of Estimates Including Parameter Uncertainty

200 Residuals for Each Group Size



Standard Deviation of  
Residual:

0.014 for 10 subjects

0.006 for 50 subjects

0.005 for 100 subjects

0.003 for >200 subjects

Suggested 95%

Confidence Intervals:

$\pm 0.03$  for 10 subjects

$\pm 0.02$  for 50 subjects

$\pm 0.01$  for 100 subjects

$\pm 0.006$  for >200 subjects

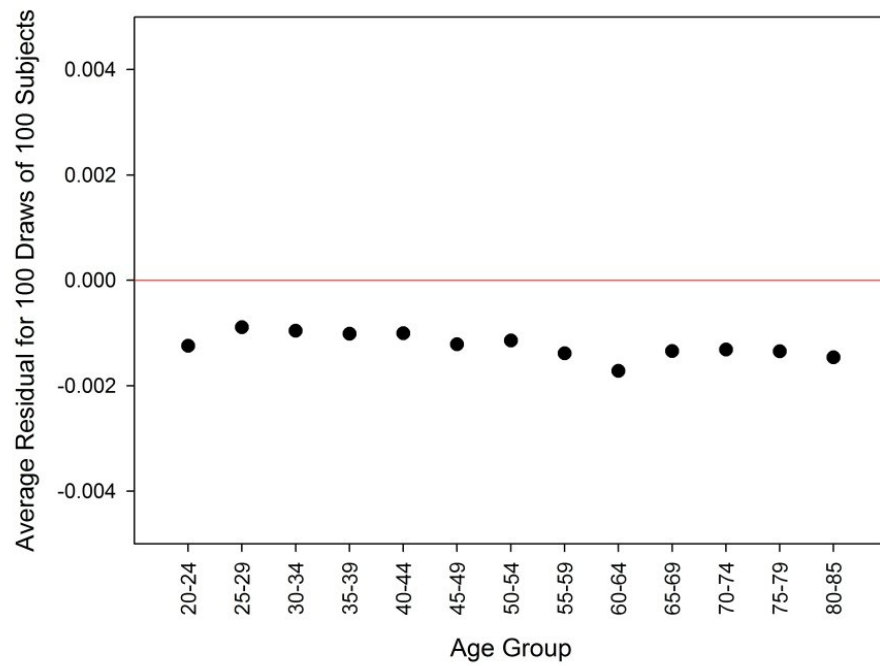
$$\text{SF-6D} = .761 \pm 0.02$$



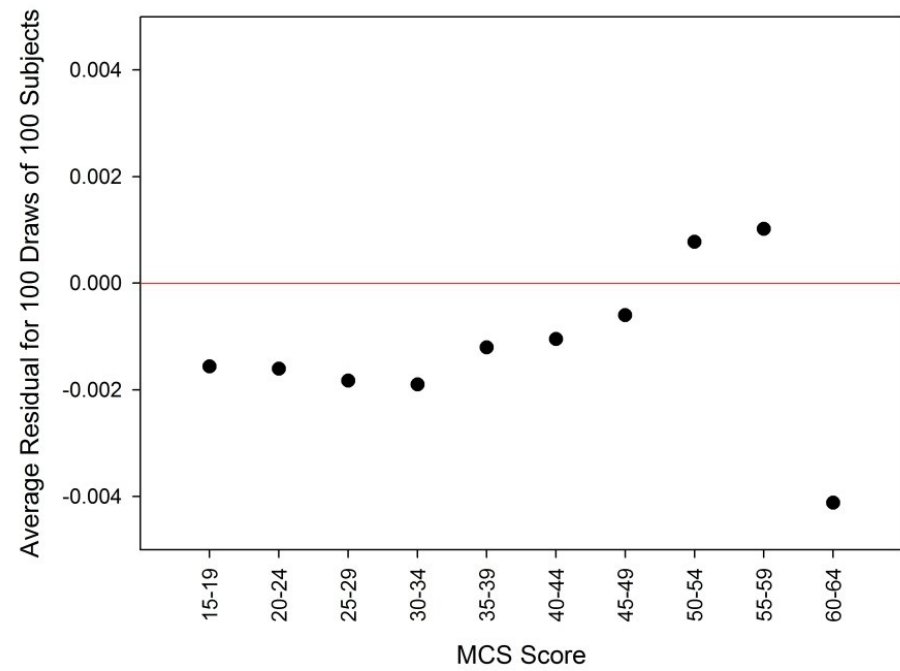
Dept. of Population Health Sciences  
University of Wisconsin-Madison

# Model Bias

Bias By Age Group



Bias by MCS Score





Dept. of Population Health Sciences  
University of Wisconsin-Madison

## Using this Algorithm with Version 2 of the SF-12

- ❑ MCS and PCS scores are dependent on the normative population values used:
  - MEPS used 1990 norms for version 1
  - MEPS used 1998 norms for version 2
- ❑ Using our algorithm with MCS and PCS scores from version 2 results in a bias of 0.0185
- ❑ Estimate error is the same as version 1
- ❑ We suggest predicting SF-6D scores from version 2 with:  
$$\text{SF-6D} = -.01904 - .0022 * \text{Female} + .00014 * \text{Age} + .0081 * \text{MCS} + .0081 * \text{PCS}$$



Dept. of Population Health Sciences  
University of Wisconsin-Madison

# Conclusions

- ❑ If individual level data are available, SF-6D scores should be directly computed.
- ❑ Commonly reported demographics and SF-12 variables can be used to estimate an average SF-6D score without accessing individual level data.
- ❑ The confidence interval around these point estimates is dependent on sample size.



Dept. of Population Health Sciences  
University of Wisconsin-Madison

# Thank you

We look forward to your comments and questions.

[jehanmer@wisc.edu](mailto:jehanmer@wisc.edu)



Dept. of Population Health Sciences  
University of Wisconsin-Madison

## Areas of Nonsense Predictions

- ❑ The lowest score from the SF-6D scoring algorithm is 0.345. Values lower than 0.345 could be generated with:
  - $PCS + MCS < 5$ . Scores this low are not observed in MEPS
- ❑ The highest score from the SF-6D scoring algorithm is 1.0. Values exceeding 1.0 could be generated with:
  - Average age of 50,  $PCS + MCS > 115$
  - Average age of 60,  $PCS + MCS > 113$
  - Average age of 70,  $PCS + MCS > 111$