

## The Medical Outcomes Study 36-Item Short-Form Health Survey

The Medical Outcomes Study 36-Item Short-Form Health Survey (MOS SF-36 Health Survey) is a measure of health status designed for use in clinical practice, research, health policy evaluations, and general population surveys. It includes eight scales that assess the following general health concepts: physical functioning, role limitations due to physical health problems (role physical<sup>®</sup>), bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems (role emotional<sup>®</sup>), and mental health (McHorney et al. 1994).

The MHSIP Mental Health Report Card includes two of the instrument's scales: ASocial Functioning<sup>®</sup> and Arole Emotional<sup>®</sup>. The ASocial Functioning<sup>®</sup> scale measures the degree to which individual's emotional or physical problems disrupt his/her normal social activities. There are five response choices for the questions of this scale that range from Anot at all<sup>®</sup> to Aextremely<sup>®</sup> for one question and from Aall of the time<sup>®</sup> to Anone of the time<sup>®</sup> for another question. The ARole Emotional<sup>®</sup> scale measures the degree to which individual's emotional problems interfere with his/her work or other daily activities. When assessed on this scale, subjects are asked to choose between Ayes<sup>®</sup> and Ano<sup>®</sup> answers.

Table 5 summarizes the evidence of the ASocial Functioning<sup>®</sup> and ARole Emotional<sup>®</sup> scales' validity and reliability. The data analyzed in the table came from the Medical Outcomes Study (MOS) conducted in 1986-1987. The number of persons (N) included in each analysis varied and is provided in the table. Six clinical groups known to differ in severity of medical or psychiatric conditions were defined in order to complete some of the scales' validity tests. These groups were: Group 1, minor (uncomplicated) chronic medical conditions only; Group 2, serious chronic medical conditions only; Group 3, psychiatric conditions only; and Group 4, both serious medical and psychiatric conditions. In addition, for some tests Group 3 was divided into a clinical depression subgroup, which included persons with current unipolar affective disorder, and a symptomatic depression subgroup, which included persons with serious depressive symptoms in the absence of a disorder.

The reliability of the ASocial Functioning<sup>®</sup> and ARole Emotional<sup>®</sup> scales as verified by means of internal consistency tests (table 5, rows 1-7). Their validity was concluded based on the scales' association with other health status measures and on their ability to discriminate between groups known to be clinically different. All but two F statistics calculated for validity tests were significant<sup>1</sup> (table 5, rows 21-23; 32-38). Their absolute values, however, were essentially different. Therefore, Arelative validity<sup>®</sup> (RV) estimates were used in order to provide a point of reference for the obtained validity indicators. Relative validity can be defined as a ratio of the validity coefficients of the scales' in question relative to the validity coefficients of some standard scales. Here, validity coefficients of the ASocial Functioning<sup>®</sup> and ARole Emotional<sup>®</sup> scales were compared to the validity coefficients of other SF-36 scales (table 5, rows 19-20; 24-28; 39-45), the MOS long-form and global scales (Stewart and Ware 1992) and Dartmouth COOP charts (Nelson et al. 1987) (table 5, rows 46-47) as well as to

---

<sup>1</sup>In table 5, F statistics that were not significant even at .05 level are marked with [Uns]

other validity indicators of the scales themselves (table 5, rows 18; 29-30; 48-53). The highest relevant coefficient (table 5, rows 18-20; 24-30; 39-45; 48-53) or a mean of all relevant coefficients (table 5, rows 46-47) were used for comparison. For example, relations of the ASocial Functioning® and ARole Emotional® scales with the SF-36 global mental and physical health dimensions were compared to those of the survey's AMental Health® and APhysical Functioning® scales that respectively were most associated with the global dimensions of mental and physical health (table 5, rows 19-20).

The relative validity coefficients demonstrated that both ASocial Functioning® and ARole Emotional® scales were most related to the global mental health SF-36 factor (table 5, row 18). Further, both scales were found to discriminate best between psychiatric (Group 3) and uncomplicated chronic medical (Group 1) conditions (table 5, rows 48-53). The estimates obtained for each scale's ability to discriminate between these groups were higher than the ability estimates averaged among each scale itself and the corresponding MOS long-form and global scales and Dartmouth COOP charts. In addition, both scales have shown a rather high sensitivity to the cross-sectional and longitudinal differences in the severity of clinical depression (table 5, rows 27-28). The two scales differed in that the ARole Emotional® scale was found to be associated with the global self-evaluation of changes only in mental health, whereas the ASocial Functioning® scale predicted the self reported changes in both mental and physical health as well as general health quite well (table 5, rows 24-26; 29-30).

The analysis presented in table 5 also includes the calculated proportions of respondents who obtained the lowest possible and the highest possible scores on the scales (rows 8-15). These proportions revealed rather strong ceiling effects for both scales in the aggregated sample of individuals with various diagnoses (table 5, rows 11 & 15) and in the symptomatic depression group (table 5, row 13). Floor effects were somewhat high for the ARole Emotional® scale in the clinical depression group (table 5, row 9). The presence of the ceiling and floor effects should be taken into consideration when using the ASocial Functioning® and ARole Emotional® SF-36 scales within the MHSIP Mental Health Report Card.

**Table 5**

Study	Row No.	Estimates	Scale	
			Social Functioning	Role Emotional
McHorney et al. 1994	1	Internal consistency (individual item correlations with the scale corrected for overlap; N=3,445)	.74	.63-.73
	2	Internal consistency obtained for Clinical Depression Group (N=503)	.70	.50-.69
	3	Internal consistency obtained for Symptomatic Depression Group (N=785)	.66	.61-.68
	4	Reliability (alpha; N=3,445)	.85	.83
	5	Alpha obtained for Clinical Depression Group (N=503)	.82	.77
	6	Alpha obtained for Symptomatic Depression Group (N=785)	.79	.80
	7	Homogeneity (average inter-item correlation; N=3,445)	.74	.61
	8	Floor (N=3,445)	.8%	18.1%
	9	Floor for Clinical Depression Group (N=503)	2.6%	41.6%
	10	Floor for Symptomatic Depression Group (N=785)	.8%	25.9%
	11	Ceiling (N=3,445)	46.2%	56.1%
	12	Ceiling for Clinical Depression Group (N=503)	13.2%	22.8%
	13	Ceiling for Symptomatic Depression Group (N=785)	30.1%	39.9%

McHorney et al. 1992	14	Floor (N=969)	.4%	14.1%
	15	Ceiling (N=969)	54.2%	64.2%

Study	Row No.	Estimates	Scale	
			Social Functioning	Role Emotional
<i>Construct Validity</i>				
Estimates are calculated based on McHorney et al. (1993)	16	The scale-s total variance explained by the SF-36 mental health factor	50.4%	65.6%
	17	The scale-s total variance explained by the SF-36 physical health factor	19.4%	3.6%
	18	$RV^2$ (the scale-s total variance explained by the SF-36 physical health factor; compared to that explained by the SF-36 mental health factor)	.38	.05

<sup>2</sup>RV stands for relative validity. For definition see page 28.

McHorney et al. 1993	19	RV (the scale's total variance explained by the Sf-36 mental health factor; compared to Mental Health@SF-36 scale)	.62	.81
	20	RV (the scale's total variance explained by the SF-36 physical health factor; compared to Physical Functioning@SF-36 scale)	.25	.05
Mc Horney et al. 1995	21	F statistic (self-reported changes in mental health; N=1,442)	31.3	27.2
	22	F statistic (self-reported changes in physical health N=1,513)	27.5	5.1
	23	F statistic (self-reported changes in general health N=1,454)	36.0	6.3
	24	RV global self evaluation of changes in mental health; compared to Mental Health@SF-36 scale)	.53	.46
	25	RV global self evaluation of changes in physical health; compared to Physical Functioning@SF-36 scale)	.61	.11
	26	RV global self evaluation of changes in general health; compared to General Health Perceptions@SF-36 scale)	.65	.11
	27	RV (severity of clinical depression; compared to Mental Health@SF-36 scale)	.58	.41
	28	RV (changes in severity of clinical depression; compared to Vitality@SF-36 scale)	.52	.84
Estimates are calculated based on McHorney et al. 1995	29	RV (global self-evaluation of changes in general health; compared to those in mental health)	.88	.19
	30	RV (global self-evaluation of changes in general health; compared to those in mental health)	1.15	.23

Study	Row No.	Estimates	Scale	
			Social Functioning	Role Emotional
<i>Discriminant Validity</i>				
McHorney et al. 1994	31	The scale's correlation with other SF-36 scales (N=3,445)	.20-.62	.11-.56
McHorney et al. 1993	32	F statistic (Group 3 (N=153) vs. Group 1 (N=576))	158.6	159.9
	33	F statistic (Group 4 (N=43) vs. Group 1 (N=576))	57.4	27.4
	34	F statistic (Group 4 (N=43) vs. Group 2 (N=144))	12.9	12.9
	35	F statistic (Group 4 (N=43) vs. Group 3 (N=153))	.0[Uns]	3.1[Uns]
	36	F statistic (Group 3 (N=153) vs. Group 2 (N=144))	28.7	62.7
	37	F statistic (Group 2 (N=144) vs. Group 1 (N=576))	29.9	5.8
	38	F statistic (Symptomatic (N=56) vs. Clinical (N=97) Depression Groups)	15.64	21.1
	39	RV (Group 3 vs. Group 1; compared to Mental Health; SF-36 scale)	.54	.54
	40	RV (Group 4 vs. Group 1; compared to General Health Perceptions SF-36 scale)	.61	.29
	41	RV (Group 4 vs. Group 2; compared to Mental Health SF-36 scale)	.34	.34
	42	RV (Group 4 vs. Group 3; compared to Physical Functioning SF-36 scale)	.00	.05
	43	RV (Group 3 vs. Group 2; compared to Mental Health SF-36 scale)	.20	.45
	44	RV (Group 2 vs. Group 1; compared to Physical Functioning SF-36 scale)	.35	.07

45	RV (Symptomatic (N=56) vs. Clinical (N=97) Depression; compared to Mental Health@SF-36 scale)	.32	.43
----	---	-----	-----

Study	Row No.	Estimates	Scale	
			Social Functioning	Role Emotional
McHorney et al. 1992	46	RV (Group 1 (N about 600) vs. Group 3 (N=155); compared to the corresponding MOS long-form and global scales as well as Dartmouth COOP charts)	1.04	1.31
	47	RV (Group 1 (N about 600) vs. Group 2 (N about 150); compared to the corresponding MOS long-form and global scales as well as Dartmouth COOP charts)	.93	.19
Estimates are calculated based on Mc Horney et al. 1995	48	RV (Group 4 vs. Group 1; compared to Group 3 vs. Group 1)	.36	.17
	49	RV (Group 4 vs. Group 2; compared to Group 3 vs. Group 1)	.08	.08
	50	RV (Group 4 vs. Group 3; compared to Group 3 vs. Group 1)	.00	.02
	51	RV (Group 3 vs. Group 2; compared to Group 3 vs. Group 1)	.18	.39
	52	RV (Group 2 vs. Group 1; compared to Group 3 vs. Group 1)	.19	.04
	53	RV (Symptomatic (N=56) vs. Clinical (N=97) Depression; compared to Group 3 vs. Group 1)	.10	.13