



University of
South Australia

Comparison of measures

Julie Ratcliffe

- Brazier et al (2007) report on the results of 24 studies published in the last 10 years that compared two or more of the instruments empirically: wide diversity of interventions and populations
- General agreement found to be poor to moderate (0.3-0.5) as measured by the intraclass correlation coefficient
- Differences in mean scores for EQ-5D, SF-6D and HUI3 comparisons found to be around 0.05 (Brazier et al 2004; Hatoum et al 2004; O'Brien et al 2004) but mask large differences in distribution of scores e.g. Brazier et al found that some respondents with EQ-5D values of 1.00 had corresponding SF-6D scores as low as 0.56!



University of
South Australia

Why do these differences exist?

- The scales do not cover the same aspects of health
- Differences in whether e.g. physical functioning is described in terms of capacity (e.g. HUI) or actual behaviour and performance (e.g. QWB)
- HUI3 covers some symptoms and a number of key impairments e.g. vision and hearing loss but unlike SF-6D and EQ-5D does not examine role or social function specifically
- The 15D and AQoL cover the most number of dimensions and levels of any generic preference based measure: defining 31 billion and 16.8 million health states respectively



University of
South Australia

Sensitivity of dimensions

- The instruments differ considerably in the number of levels in each dimension and the range of severity covered
- Floor effect in SF-6D when compared to e.g. EQ-5D appears to be more pronounced for physical functioning, role limitations and social functioning (Longworth and Bryan 2003; Brazier et al 2004)
- Conditions focused more on physical health problems: leg ulcers, osteoarthritis, rather than pain: lower back pain and irritable bowel syndrome
- Ceiling effect in EQ-5D relative to SF-6D. Brazier et al (2004) found that majority of people in state 11111 according to EQ-5D reported some problems according to SF-6D



University of
South Australia

Valuation methods

- HUI2, HUI3 and SD-6D use standard gamble (SG)
- AQL and EQ-5D use time trade off (TTO)
- 15D and QWB use visual analogue scale (VAS)
- Evidence suggests SG and TTO tend to give higher values for same health state than VAS (e.g. Torrance et al 1976: Bombardier et al, 1982: Bass et al 1994) and SG values tend to be higher than TTO values for same health state although Tsuchiya et al (2006) found a 'cross over' effect
- Impact of variants of valuation methods can also have a large impact e.g. Ratcliffe et al (2007) found major differences between values based upon titration vs ping pong methods for SG



University of
South Australia

Which measure should be used?

- Practicality / Reliability
- Descriptive validity e.g AQoL and HUI3 may be more sensitive for visual impairment conditions than EQ-5D or SF-6D
- Valuation: QWB and 15D valued using a non choice based method (VAS). SF-6D uses SG, AQoL and EQ-5D use TTO. Source of values – country specific
- Empirical validity: Comparative evidence is limited: EQ-5D found to be more responsive than HUI3 in musculoskeletal disease and liver disease (Spady and Suarez-Almazo, 2001; Longworth and Bryan 2003) but less responsive in rehabilitation patients (Kohlman et al 2005)