# SF-6D Population Norms for the Hong Kong Chinese General Population

Authors: Carlos K.H. Wong[1] (ORCID: 0000-0002-6895-6071), Brendan Mulhern[2] (ORCID: 0000-0003-3656-8063), Garvin H.L. Cheng[1], Cindy L.K. Lam[1]

[1] Department of Family Medicine and Primary Care, The University of Hong Kong

[2] Centre for Health Economics Research and Evaluation, University of Technology Sydney

**Corresponding Author:** Carlos K.H. Wong, Department of Family Medicine and Primary Care, The University of Hong Kong. Address: Rm 1-01, 1/F, Jockey Club Building for Interdisciplinary Research, 5 Sassoon Road, Pokfulam, Hong Kong. Tel: (+852) 2831-5055 Fax: (+852) 2814-7475. Email: carlosho@hku.hk

**Keywords**: Health-related quality of life; population norm; normative value; SF-6D; health utility; Chinese; Hong Kong

**Running Title**: Hong Kong Population Norms for SF-6D

**Conflict of interest statement**: Financial support for this study was provided by Health Services Research Committee (HSRC#711026), Hong Kong SAR. The funding agreement ensured the authors independence in designing the study, interpreting the data, writing, and publishing the report.

**Informed consent**: Informed consent was obtained from all individual participants included in the study.

**Competing interest:** None declared

**Abstract**

**Purpose:** To estimate population norms for the SF-6D health preference (utility) scores derived from the MOS SF-36 version 1 (SF-36v1), SF-36 version 2 (SF-36v2), and SF-12v2) Health Surveys collected from a representative adult sample in Hong Kong, and to assess differences in SF-6D scores across sociodemographic subgroups.

**Methods:** A random telephone survey of 2410 Chinese adults was conducted. All respondents completed questionnaires on socio-demographics and presence of chronic diseases (hypertension, diabetes, chronic rheumatism, chronic lung diseases, stroke, and mental illness), and the short-form 36-item Health Survey (SF-36) version 1 and selected items of the SF-36v2 that were different from those of SF-36v1. Responses of short-form 12-item Health Survey (SF-12) were extracted from responses of the SF-36 items. SF-6D health utility scores were derived from SF-36 version 1 ($SF\text{-}6D_{SF\text{-}36v1}$), SF-36 version 2 ($SF\text{-}6D_{SF\text{-}36v2}$), and SF-12 version 2 ($SF\text{-}6D_{SF\text{-}12v2}$) using Hong Kong SF-6D value set.

**Results**: Population norms of $SF\text{-}6D_{SF\text{-}36v1}$, $SF\text{-}6D_{SF\text{-}36v2}$ and $SF\text{-}6D_{SF\text{-}12v2}$ for the Hong Kong Chinese were 0.7947 (±0.0048), 0.7862 (±0.0049) and 0.8147 (±0.0050), respectively. Three SF-6D scores were highly correlated (0.861-0.954), and had a high degree of reliability and absolute agreement. Males had higher health utility scores ($SF\text{-}6D_{SF\text{-}36v1}$: 0.0025; $SF\text{-}6D_{SF\text{-}36v2}$: 0.025; $SF\text{-}6D_{SF\text{-}12v2}$:0.018) but reported less problems in all the dimensions than women. Respondents with a higher number of chronic diseases had lower SF-6D scores. Among all respondents with one or more chronic diseases, those with hypertension scored the highest whereby those mental illness scored the least.

**Conclusions**: The SF-6D utility scores derived from different SF-36 or SF-12 Health Surveys were different. The population norms based on these three Health Surveys enable the normative comparisons of health utility scores from specific population or patient groups, and provides estimates of age-gender adjusted health utility scores for health economic evaluations.

**Manuscript Text**

## Introduction

Cost-utility analysis is used to measure the cost effectiveness of emerging healthcare technology when compared with conventional technology (1). It uses the quality-adjusted life year (QALY) as the outcome, accounting for the disease burden (morbidity) and mortality over the assessment period (1). Morbidity refers to preference weighting in the form of a health utility score measuring the quality of life for an individual, where the values of 0 and 1 are interpreted as dead and full health, respectively. Generic multi-attribute utility instruments, such as the Short-form 6-dimension (SF-6D) (2), EuroQol 5-dimension (EQ-5D) (3) and Health Utility Index (HUI) (4), measure different dimensions of the quality of life using a health state classification system (5). The estimation of health utility scores is based on country-specific scoring algorithms indicating the preference weights of each health state described by the classification system (5).

Assessing health-related quality of life and health utility scores in population health surveys enables the estimation of population norms or reference values to allow for comparisons of different population and clinical groups to inform health care and policy (6). Normative comparisons, which compare the health utility scores of demographic subgroups (e.g. age, gender and socio-economic status), identify subgroups which deviate from normative values and the burden of certain diseases (7). Such information provides insight into health issues at the local population level and may inform policy issues (8). Local normative data facilitate not only within-country normative comparisons but also comparisons between countries, regions and ethnic populations. Health utility scores stratified by age also provide “age-appropriate” utility scores for comparison across the general population and people with a certain condition (9).

The wider use of multi-attribute utility instruments in part relies on the availability of respective scoring algorithms for certain country and ethnic populations. Although the EQ-5D is the most widely used generic multi-attribute utility instrument, a validated scoring algorithm or value set is not yet available for the Hong Kong Chinese population. With the development and validation of the SF-6D scoring algorithm in Hong Kong (10, 11), the derivation of SF-6D health utility norms in the Hong Kong Chinese population using available data sources containing SF-36 and SF-12 Health Surveys would inform health care and policy (8). The SF-6D utility score serves as preference weighting input to QALY outcome in economic evaluation. To our best knowledge, SF-6D population norms have been published for the general population of the UK (8), Australia (12), the US (13), Brazil (14), Chile (15), Portugal (16) and Japan (17). However, SF-6D population norms are not yet available for the Hong Kong general population. Furthermore, none of the aforementioned population norm studies compared SF-6D scores derived from item responses of different Health Surveys. The main aim of this study is to estimate Hong Kong population norms for the SF-6D scores derived from the SF-36 and SF-12 Health Surveys. Specific objectives are to compare the SF-6D scores derived from item responses of three Health Surveys, and to assess differences in health utility scores across demographic subgroups.

## Method

*Study design, and subject and study instruments*

Data analysed in this study came from a cross-sectional random telephone survey of a representative Hong Kong Chinese adult population. Household telephone numbers were randomly selected from residential telephone directories with 90% coverage of Chinese residential telephone numbers in Hong Kong from June to September 1998. Non-Chinese households, commercial numbers, telephone numbers that were unanswered after three attempts, and households without any Chinese adults were not included in the survey. Details of the sampling and survey methods are available in earlier papers (18, 19). All respondents completed questionnaires on socio-demographic status and the occurrence of chronic disease (hypertension, diabetes, chronic rheumatism, chronic lung disease, stroke and mental illness) using the short-form 36-item Health Survey version 1 (SF-36v1) and items number 4 and 5 from the short-form 36-item Health Survey version 2 (SF-36v2). The item responses for the SF-36v2 and SF-12v2 Health Surveys were derived from the item responses of the SF-36v1 Health Survey. Chronic disease included the following: hypertension, diabetes mellitus, arthritis or other chronic rheumatism, chronic lung disease (asthma, emphysema, chronic bronchitis or other chronic lung disease), stroke and mental illness (depression, anxiety, neurasthenia, psychosis or other mental illness). We counted the number of co-existing chronic diseases for each respondent.

<u>The Short-Form 36-item (SF-36) Health Survey</u>

The SF-36 Health Survey is a popular generic HRQOL measure. The Chinese (Hong Kong) SF-36v1 and SF-36v2 have been translated and validated for the Chinese adult population in Hong Kong (18). The SF-36v2 Health Survey is a newer version of version 1, with an

improvement in item layout, simpler instructions, and the changes in response options for items in the energy, fatigue, mental health scales (from level 6 to level 5), and role physical and role emotional scales (from level 2 to level 5) (20). The SF-36 has eight scales measuring eight domains of HRQOL, namely physical functioning (PF); role-physical (RP); role-emotional (RE); bodily pain (BP); general health (GH); vitality (VT); social functioning (SF); and mental health (MH). Each scale has a score ranging from 0 to 100 with a higher score indicating a better HRQOL.

The Short-Form 12-Item (SF-12) Health Survey

The SF-12 Health Survey is an abbreviated version of the SF-36 Health Survey. The SF-12 includes two physical functioning items about moderate activities and climbing stairs (11 in SF-36), two role physical items (4 in SF-36), one pain item about the extent of interference with normal work activities due to pain (2 in SF-36), one general health item about the general rating of health (5 in SF-36), one vitality item (4 in SF-36), one social functioning item (2 in SF-36), two role emotional items (3 in SF-36), two mental health items about depression and psychological well-being (5 in SF-36). The Chinese (Hong Kong) SF-12v2 has been translated and validated for the Chinese population in Hong Kong (1f9). The SF-12v2 summary scores explained 85% and 87% of the total variances of the Chinese (HK) SF-36 PCS and MCS, respectively. The SF-12 measures the same domains as does SF-36, on a theoretical scale ranging from 0 to 100. The SF-12 is valid and reliable, and considered as a substitute for the SF-36(19). A higher score implies a better HRQOL.

The Chinese (HK) SF-36v2 Health Survey data were extracted from the responses to item numbers 1, 3, 6, 7, 8, 9, 10 and 11 of version 1, which were the same as those of version 2, and the responses to the SF-36v2 Health Survey item numbers 4 and 5. The Chinese (HK) SF-12v2 Health Survey data were extracted from the responses to item numbers 3, 8, 9 and

10 of SF-36v1 and item numbers 4 and 5 of the SF-36v2. The response value "a good bit of the time" for item 9 was recorded randomly to the adjoining values of "most of the time" or "some of the time" for the calculation of the SF-36v2 and SF-12v2 vitality and mental health scale scores. Item responses recoding and scale score calculation follow the standard methods mentioned in the SF-36v2 Health Survey manual and SF-12v2 Health Survey manual.

*Outcome measures*

Short-form 6-dimension (SF-6D)

The SF-6D instrument, developed by Brazier et al. (2), has six items with four to six response levels for each item. Each dimension (physical functioning, role limitation, social functioning, pain, mental health and vitality) is represented by one item. Response level 1 indicates no problem in that dimension, whereas levels with larger number indicate a more serious problem in that dimension.

The Hong Kong SF-6D scoring algorithm (10, 11) was derived by the standard gamble valuation method. The theoretical range of the SF-6D utility score was from 1 for full health to 0.315 for the worst possible health state according to the Hong Kong Chinese population-specific scoring algorithm (10, 11). In this study, the SF-6D health utility score was not converted from the raw response levels of the SF-6D instrument. Instead, response levels of six dimensions in the SF-6D were mapped from the available item responses of the short-form Health Surveys SF-36v1, SF-36v2 and SF-12v2. We denote $\text{SF-6D}_{\text{SF-36v1}}$, $\text{SF-6D}_{\text{SF-36v2}}$ and $\text{SF-6D}_{\text{SF-12v2}}$ as the SF-6D score derived from the SF-36v1, SF-36v2 and SF-12v2 Health Survey, respectively. Scoring algorithms entails using the SF-6D response levels mapped from three Health Surveys and corresponding utility decrements due to more severe problems in each dimension (see *Supplemental Material*).

*Statistical Analysis*

Means and their corresponding 95% confidence intervals (CIs) of SF-6D scores by gender, by age group and by number of chronic diseases were calculated to show descriptive statistics of three scores ($SF\text{-}6D_{SF\text{-}36v1}$, $SF\text{-}6D_{SF\text{-}36v2}$ and $SF\text{-}6D_{SF\text{-}12v2}$). Independent t-test and analysis of variance test were performed to assess the differences in SF-6D scores between socio-demographic groups. Chi-square tests were performed to test the differences in ceiling effect of the scores between men and women. Analysis was performed to investigate the differences in the SF-6D scores with respect to different levels in age groups and numbers of chronic diseases. Variance-weighted least squares regression was used to test the trend of an increasing number of chronic diseases for the three scores.

To test the reliability across the three SF-6D health utility scores, a two-way mixed model was constructed to calculate an intra-class correlation coefficient (ICC). ICC ranges from 0 to 1, where 0 indicates no agreement and 1 indicates perfect agreement. The Bland-Altman plots were established to check absolute agreement of different health utility scores. In the Bland-Altman plot, the difference in the SF-6D scores between two Health Surveys was put as the y-axis while the mean SF-6D scores of two respective Health Surveys was put as the x-axis. The mean difference and their 95% CI bounds were shown in the plot.

The response options of SF-12 and SF-36 were mapped onto response levels in each dimension of the SF-6D. The proportion of respondents with no problem in each SF-6D dimension was reported.

All data analyses were conducted with the SPSS for Windows 24.0 programme. Statistically significant levels were set at p-value less than 0.05.

## Results

In total, 2,410 Chinese adults in Hong Kong completed the interview, a response rate of 84.4% (2,410 out of 2,857). Of these, 47.8% were men and the mean age of respondents was 42.9 years (standard deviation of 17.3; ages ranging from 18–88 years). The socio-demographic characteristics of the study samples were comparable to those of the adult general population in 1996 (49.5% men; mean age, 42.3 years) (21). The differences in the age and gender distribution were not statistically significant (all p-values >0.05). When compared with the current general population in 2016 (45.1% men; mean age, 47.9 years) (22), the mean age of our study samples was about 5 years younger and more men were included.

Table 1 shows the descriptive statistics of the SF-6D scores stratified by gender and age. The SF-6D$_{SF-36v1}$, SF-6D$_{SF-36v2}$ and SF-6D$_{SF-12v2}$ for the sample were 0.7947 (±0.0048), 0.7862 (±0.0049) and 0.8147 (±0.0050), respectively. Men reported higher mean SF-6D scores than women (all P<0.001). The SF-6D$_{SF-36v1}$, SF-6D$_{SF-36v2}$ and SF-6D$_{SF-12v2}$ scores of men were 0.8078 (±0.0067), 0.7992 (±0.0068) and 0.8242 (±0.0069), respectively; those of women were 0.7828 (±0.0068), 0.7743 (±0.0070) and 0.8061 (±0.0072), respectively. The ceiling effects of SF-6D$_{SF-36v1}$, SF-6D$_{SF-36v2}$ and SF-6D$_{SF-12v2}$ among men were 2.34%, 2.34% and 3.91%, while those of women were 0.95%, 0.95% and 2.94%. The p-values in testing the ceiling effect were 0.007, 0.007 and 0.192, respectively. The ceiling effect in men was therefore greater than in women for SF-6D$_{SF-36v1}$ and SF-6D$_{SF-36v2,}$ while the ceiling effect in men was insignificantly greater than in women for SF-6D$_{SF-12v2}$. The SF-6D$_{SF-36v1}$ and SF-6D$_{SF-36v2}$ scores of respondents aged 18 to 60 years were above 0.78, but the scores declined with age from age 60. Respondents aged 21 to 30 years scored the highest for SF-6D$_{SF-36v1}$ and SF-6D$_{SF-36v2}$. For SF-6D$_{SF-12v2}$, “51 to 60” scored the most, after which the score started

to decline. The oldest age group of "81 to 90" scored the least in the three scores.

The frequency distributions of the three SF-6D scores were plotted in a histogram (Figure 1). The ICC coefficients of three SF-6D scores were higher than 0.9, meaning that the three Health Surveys were highly consistent in calculating the SF-6D utility scores. Figure 2 depicts the Bland-Altman plots for the SF-6D utility scores. The difference for $SF\text{-}6D_{SF\text{-}36v1}$ and $SF\text{-}6D_{SF\text{-}36v2}$ was 0.0085; the difference for $SF\text{-}6D_{SF\text{-}36v2}$ and $SF\text{-}6D_{SF\text{-}12v2}$ was 0.0285; and the difference for $SF\text{-}6D_{SF\text{-}36v1}$ and $SF\text{-}6D_{SF\text{-}12v2}$ was 0.0200. The most outliers lying outside the 95% CIs were observed in the plot comparing the agreement between $SF\text{-}6D_{SF\text{-}36v1}$ and $SF\text{-}6D_{SF\text{-}12v2}$. Most of the outliers occurred at a high SF-6D score ($\geq$0.7). Figure 3 shows the scatter plots of the SF-6D scores. All scatter plots reveal a positive linear relationship among the three SF-6D scores. The Pearson correlation coefficients ranged from 0.861 to 0.954, implying extremely high correlations between the SF-6D scores.

Table 2 shows the scores stratified by the type of chronic disease. Respondents who were free of chronic disease had the highest health utility scores. Among all chronic diseases considered in this survey, respondents with hypertension scored the highest whereby respondents with mental illness scored the lowest. Figure 4 shows that the SF-6D scores decrease with an increasing number of chronic diseases. The $SF\text{-}6D_{SF\text{-}36v1}$, $SF\text{-}6D_{SF\text{-}36v2}$ and $SF\text{-}6D_{SF\text{-}12v2}$ of respondents with five to six chronic disease were 0.5983, 0.5952 and 0.6298, which is far below the scores (0.7947, 0.7862 and 0.8147, respectively) for the whole sample. For the variance-weighted least square regression, the coefficients for the number of chronic diseases for $SF\text{-}6D_{SF\text{-}36v1}$, $SF\text{-}6D_{SF\text{-}36v2}$ and $SF\text{-}6D_{SF\text{-}12v2}$ were -0.048, -0.046 and -0.038, respectively. The p-value of testing the coefficients equal 0 were all $< 0.05$, indicating that respondents suffering from a greater number of chronic diseases would result in significantly lower scores for $SF\text{-}6D_{SF\text{-}36v1}$, $SF\text{-}6D_{SF\text{-}36v2}$ and $SF\text{-}6D_{SF\text{-}12v2}$.

Table 3 summarizes the distribution of respondents reporting no problem in the six dimensions mapped from raw responses of the three Health Surveys stratified by gender. Men in general reported more "no problems" in all the dimensions than women. Conversely, women consistently reported more "no problems" in the following scenarios: role functioning dimension in respondents aged 21 to 40 years, social functioning dimension in respondents aged 21 to 30 years, mental health dimension in respondents aged 51 to 60 years and vitality dimension in respondents aged 31 to 40 years. Table 4 demonstrates the overall distribution of levels in each dimension mapped from raw item responses of the three Health Surveys. At least half the respondents reported no problem in physical functioning, role limitation, social functioning and pain; only ≤30% of respondents reported no problem in mental health and only 7.1% reported no problem in vitality.

**Discussion**

Published population norm studies have reported the SF-6D scores derived from either the SF-12 (8, 15) or the SF-36 Health Survey (12, 13, 16, 17), with the exception of the SF-6D scores computed by raw SF-6D responses in a Brazilian population norm study (14). The investigation of the Hong Kong population norms of SF-6D scores derived from item responses of three Health Surveys (the SF-36v1, SF-36v2 and SF-12v2) is the strength of this population norm study. The key finding of this study is the identification of demographic subgroups to inform health care and policy. In line with population norm studies all over the world, men had higher mean SF-6D scores than women, whereas the respondents with greater chronic disease counts had impaired SF-6D scores compared with those with lesser counts. Interestingly, the association between SF-6D scores and age groups was inconsistent across the population norm studies. The mean SF-6D scores decreased with increasing age in

the UK (8), Brazilian (14), Portuguese(16) and Australian (12) general populations. The SF-6D population norm for the Hong Kong general population was similar to that for another Asian country; a decreasing trend in age-specific SF-6D scores, with slight improvement at the age group 51 to 60 years, was also observed in the Japanese general population (17). Such differences in SF-6D scores derived from the three Health Surveys may influence the incremental cost-effectiveness ratio and policy decision making. While applying age-specific health utility for extrapolation of QALYs in health economic evaluation, respondents in older age groups had greater SF-6D scores derived from SF-12 than the SF-6D scores derived from SF-36. The use of $SF\text{-}6D_{SF\text{-}12v2}$ leads to a possible decrease in total QALYs of interventions for elderly individuals when compared with the use of a SF-36 derived SF-6D score. Varying the use of SF-6D scores in further cost-effectiveness analyses is warranted.

Regardless of age, gender and number of chronic disease stratums, the $SF\text{-}6D_{SF\text{-}12v2}$ scores were consistently higher than the $SF\text{-}6D_{SF\text{-}36v1}$ and $SF\text{-}6D_{SF\text{-}36v2}$ scores. One plausible explanation is the difference in the SF-6D preference weightings on which the response level in each dimension was mapped from the short-form Health Survey. From the scoring algorithm (see *Supplemental Material*), the most severe problems in the physical functioning dimension induced the greatest SF-6D utility decrement among the six dimensions. The utility decrement due to the most severe problems in the physical function dimension was larger (-0.178) in the $SF\text{-}6D_{SF\text{-}36v1}$ and $SF\text{-}6D_{SF\text{-}36v2}$ scoring algorithms than (-0.129) in the $SF\text{-}6D_{SF\text{-}12v2}$ scoring algorithm. Likewise, the utility decrement due to the most severe problems in the pain dimension was also larger (-0.100) in $SF\text{-}6D_{SF\text{-}36v1}$ and $SF\text{-}6D_{SF\text{-}36v2}$ scoring algorithms than (-0.082) in the $SF\text{-}6D_{SF\text{-}12v2}$ scoring algorithm. As such, the elicited SF-6D score derived from the SF-12v2 Health Survey is in general greater than that derived from the SF-36v1 and SF-36v2 Health Surveys.

Several limitations in this study should be acknowledged. First, this population telephone survey was conducted two decades ago. The demographics of the 1996 general population (21) were younger (mean 42.3 vs 47.9 years) and included a higher proportion of men (49.5% vs 45.1%) than those in the 2016 general population (22). However, the demographics of our study samples were comparable to those of general population over the sampling period. When comparing the SF-6D data collected in recent surveys or clinical trials with this SF-6D normative data, there is concern about the appropriateness of using this data as control groups for fair comparisons as the SF-6D population norms may be time-varying. Secondly, the small sample size (n=38) in the oldest age group (80–90 years) led to a large variation in estimates of the SF-6D score for that group. Indeed, the small number of samples from this population telephone survey reflected the actual proportion of older people in Hong Kong's general population. Likewise, as the number of samples for some health conditions, such as stroke (n=21), heart disease (n=94) and mental illness (n=94), were low, normative comparisons with those conditions should be interpreted with caution. These aforementioned health conditions in our study samples were under-represented while the prevalence of those health conditions in the Hong Kong population (23, 24) were higher than that in our study samples. Moreover, the three SF-6D scores had originated from the SF-36v1 Health Survey rather than three independently administered questionnaires. The actual responses to the SF-12v2 and SF-36v2 Health Surveys might then differ from those responses converted from SF-36v1 in the current secondary analysis. Finally, the use of telephone-administered data might yield biased responses in comparison with data obtained from a completely confidential and anonymous method. Respondents who were willing to take part in a telephone-administered survey limited the representativeness of the study samples in terms of health conditions and unobservable factors, despite there being a comparable socio-demographic distribution between study sample and general population.

## Conclusion

Results from this representative population telephone survey provided the SF-6D population norms for the Hong Kong Chinese population. Population norms for the SF-6D scores are derived from different short-form Health Surveys. Such population norms enable the normative comparisons of health utility scores with certain diseases, and present estimates of age–gender adjusted health utility scores for health economic evaluations. As with normative studies in other countries, evidence from the Hong Kong normative data showed that higher health utility scores were associated with respondents of male gender, younger age and absence of chronic disease.

**Compliance with Ethical Standards**

Financial support for this study was provided by Health Services Research Committee (HSRC#711026), Hong Kong SAR. The funding agreement ensured the authors independence in designing the study, interpreting the data, writing, and publishing the report.

Conflict of Interest: All authors declare that he / she has no conflict of interest.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

**Author contributions**: CKHW wrote the manuscript, researched data, contributed to statistical analysis and interpretation of results. BM contributed to interpretation of results and reviewed/edited the manuscript. GC wrote the manuscript, researched data, and contributed to statistical analysis. CLKL contributed to study design, acquisition of data and reviewed/edited the manuscript.

## References

1. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. London: NICE; 2013.
2. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of Health Economics. 2002;21(2):271-92.
3. Dolan P. Modeling valuations for EuroQol health states. Medical Care. 1997;35(11):1095-108.
4. Torrance GW, Feeny DH, Furlong WJ, Barr RD, Zhang Y, Wang Q. Multiattribute Utility Function for a Comprehensive Health Status Classification System: Health Utilities Index Mark 2. Medical Care. 1996;34(7):702-22.
5. Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. Health technology assessment (Winchester, England). 1999;3(9):i-iv, 1-164.
6. Kendall PC, Marrs-Garcia A, Nath SR, Sheldrick RC. Normative comparisons for the evaluation of clinical significance. Journal of Consulting and Clinical Psychology. 1999;67(3):285-99.
7. Williams A. Calculating the global burden of disease: time for a strategic reappraisal? Health Economics. 1999;8(1):1-8.
8. van den Berg B. SF-6D population norms. Health Economics. 2012;21(12):1508-12.
9. Clemens S, Begum N, Harper C, Whitty JA, Scuffham PA. A comparison of EQ-5D-3L population norms in Queensland, Australia, estimated using utility value sets from Australia, the UK and USA. Quality of Life Research. 2014;23(8):2375-81.
10. Lam CL, Brazier J, McGhee SM. Valuation of the SF-6D Health States Is Feasible, Acceptable, Reliable, and Valid in a Chinese Population. Value Health. 2008;11(2):295-303.
11. McGhee SM, Brazier J, Lam CL, Wong LC, Chau J, Cheung A, et al. Quality-adjusted life years: population-specific measurement of the quality component. Hong Kong medical journal. 2011;17 (Suppl 6):17-21.
12. Norman R, Church J, van den Berg B, Goodall S. Australian health-related quality of life population norms derived from the SF-6D. Australian and New Zealand Journal of Public Health. 2013;37(1):17-23.
13. Fryback DG, Dunham NC, Palta M, Hanmer J, Buechner J, Cherepanov D, et al. US norms for six generic health-related quality-of-life indexes from the National Health Measurement study. Medical Care. 2007;45(12):1162-70.
14. Ciconelli RM, Ferraz MB, Kowalski S, Pinheiro Gda R, Sato EI. Brazilian urban population norms derived from the health-related quality of life SF-6D. Quality of Life Research. 2015;24(10):2559-64.

15. Garcia-Gordillo MA, Collado-Mateo D, Olivares PR, Adsuar JC. Chilean population norms derived from the health-related quality of Life SF-6D. European Journal of Health Economics. 2017.
16. Ferreira PL, Ferreira LN, Pereira LN. SF-6D Portuguese population norms. European Journal of Health Economics. 2015;16(3):235-41.
17. Shiroiwa T, Fukuda T, Ikeda S, Igarashi A, Noto S, Saito S, et al. Japanese population norms for preference-based measures: EQ-5D-3L, EQ-5D-5L, and SF-6D. Quality of Life Research. 2016;25(3):707-19.
18. Lam CLK, Tse EYY, Gandek B, Fong DYT. The SF-36 summary scales were valid, reliable, and equivalent in a Chinese population. Journal of Clinical Epidemiology. 2005;58(8):815-22.
19. Lam ET, Lam CL, Fong DY, Huang WW. Is the SF-12 version 2 Health Survey a valid and equivalent substitute for the SF-36 version 2 Health Survey for the Chinese? Journal of evaluation in clinical practice. 2013;19(1):200-8.
20. Ware JEJ. SF-36 Health Survey Update. Spine. 2000;25(24):3130-9.
21. Government of Hong Kong SAR. 1996 Population By-census. In: Census and Statistics Department, editor. 1997.
22. Government of Hong Kong SAR. 2016 By-census Results 2017 [Available from: https://www.bycensus2016.gov.hk/en/index.html.
23. Government of Hong Kong SAR. Thematic Household Survey Report No. 63. In: Census and Statistics Department, editor. 2017.
24. Government of Hong Kong SAR. Mental Health Review Report. In: Food and Health Bureau, editor. 2017.

## Figure Legend

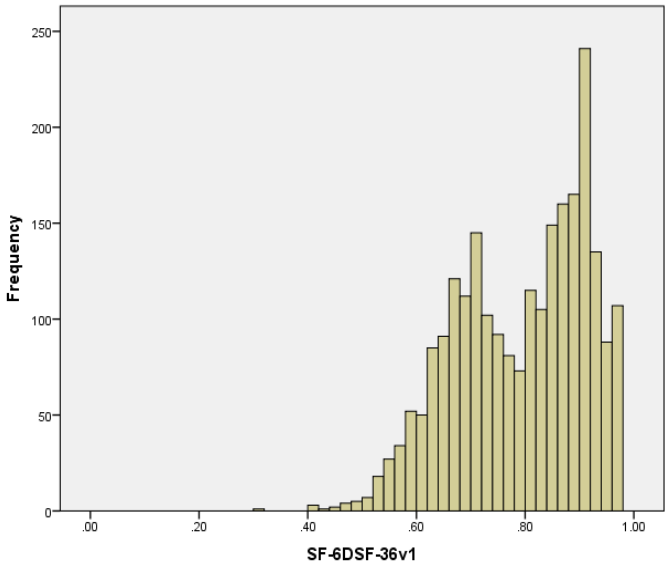
Figure 1: Frequency distribution of SF-6D health utility scores: a) $SF\text{-}6D_{SF\text{-}36v1}$; b) $SF\text{-}6D_{SF\text{-}36v2}$; c) $SF\text{-}6D_{SF\text{-}12v2}$

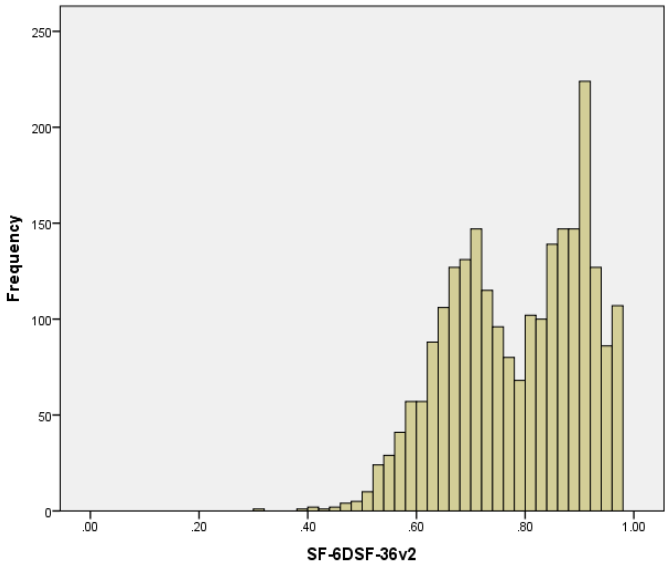
Figure 2: Bland-Altman plots for the SF-6D health utility scores: a) $SF\text{-}6D_{SF\text{-}36v1}$ vs $SF\text{-}6D_{SF\text{-}36v2}$; b) $SF\text{-}6D_{SF\text{-}36v2}$ vs $SF\text{-}6D_{SF\text{-}12v2}$; c) $SF\text{-}6D_{SF\text{-}36v1}$ vs $SF\text{-}6D_{SF\text{-}12v2}$

Figure 3: Scatter plots of the SF-6D health utility scores: a) $SF\text{-}6D_{SF\text{-}36v1}$ vs $SF\text{-}6D_{SF\text{-}36v2}$; b) $SF\text{-}6D_{SF\text{-}36v2}$ vs $SF\text{-}6D_{SF\text{-}12v2}$; c) $SF\text{-}6D_{SF\text{-}36v1}$ vs $SF\text{-}6D_{SF\text{-}12v2}$

Figure 4: SF-6D health utility scores stratified by the number of chronic diseases

Figure 1: Frequency distribution of SF-6D health utility scores: a) SF-6D$_{SF-36v1}$; b) SF-6D$_{SF-36v2}$; c) SF-6D$_{SF-12v2}$



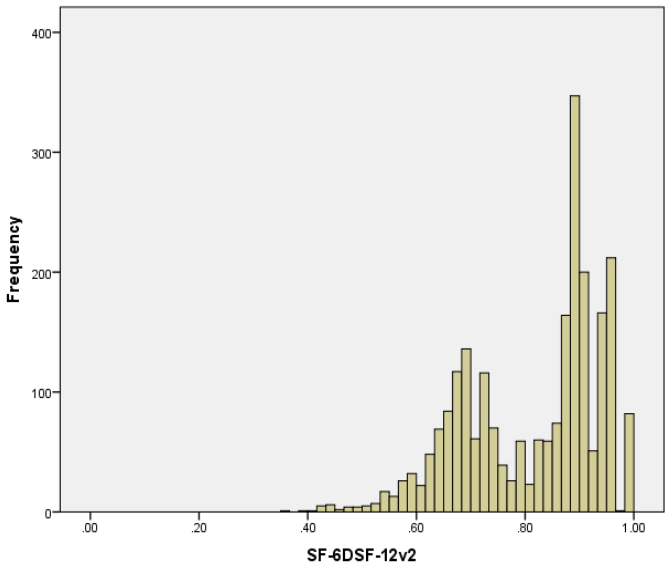
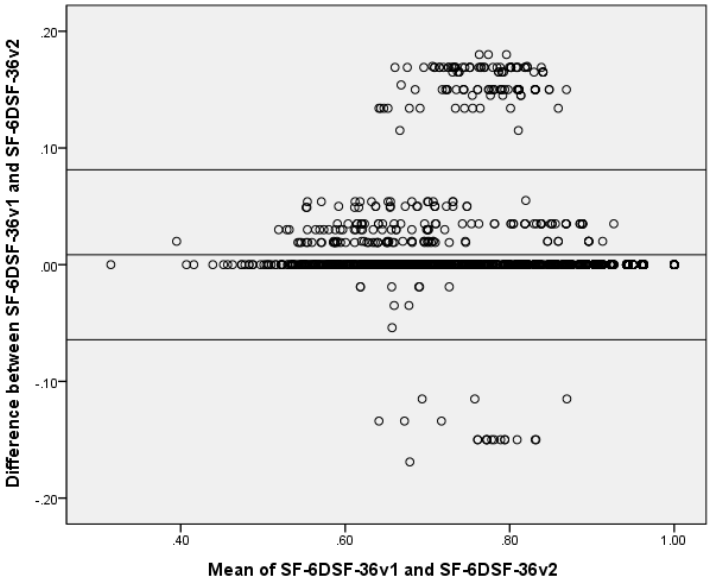
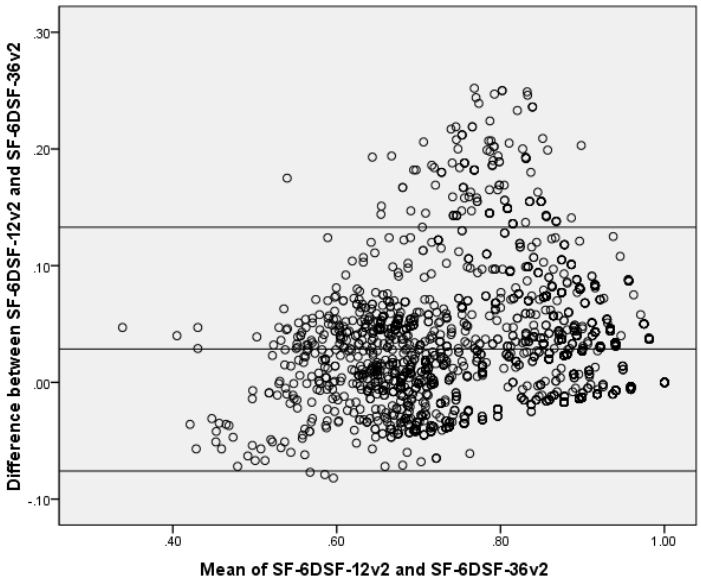



Figure 2: Bland-Altman plots for the SF-6D health utility scores: a) $SF\text{-}6D_{SF\text{-}36v1}$ vs $SF\text{-}6D_{SF\text{-}36v2}$; b) $SF\text{-}6D_{SF\text{-}36v2}$ vs $SF\text{-}6D_{SF\text{-}12v2}$; c) $SF\text{-}6D_{SF\text{-}36v1}$ vs $SF\text{-}6D_{SF\text{-}12v2}$



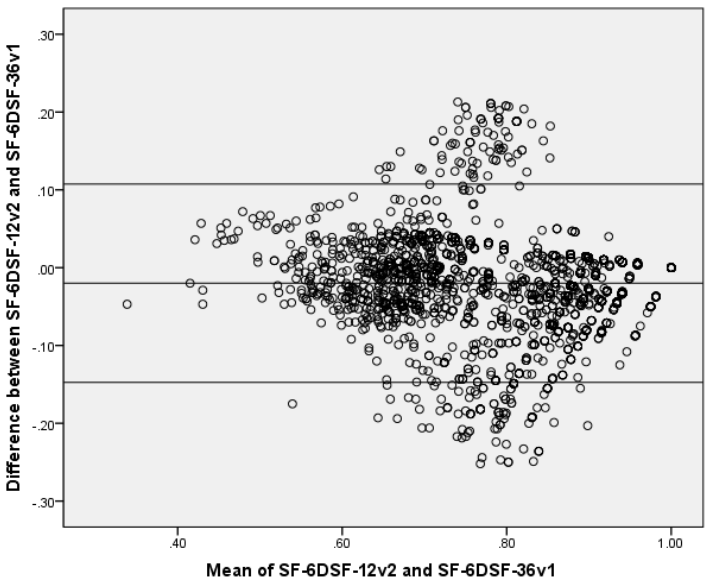
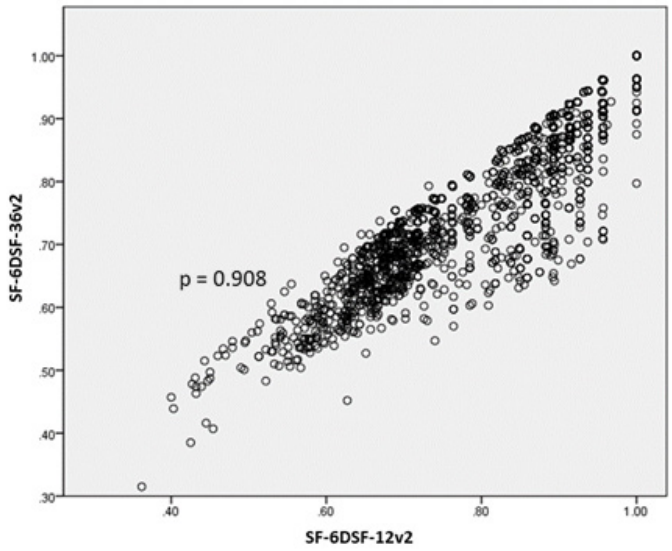
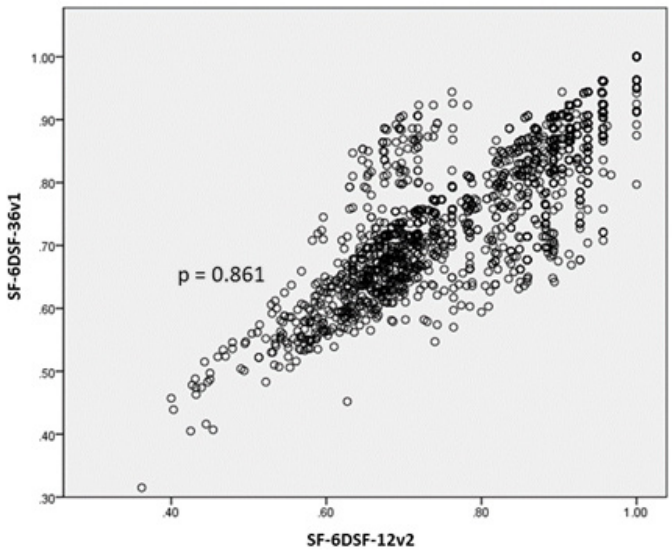



Figure 3: Scatter plots of the SF-6D health utility scores: a) $SF\text{-}6D_{SF\text{-}36v1}$ vs $SF\text{-}6D_{SF\text{-}36v2}$; b) $SF\text{-}6D_{SF\text{-}36v2}$ vs $SF\text{-}6D_{SF\text{-}12v2}$; c) $SF\text{-}6D_{SF\text{-}36v1}$ vs $SF\text{-}6D_{SF\text{-}12v2}$



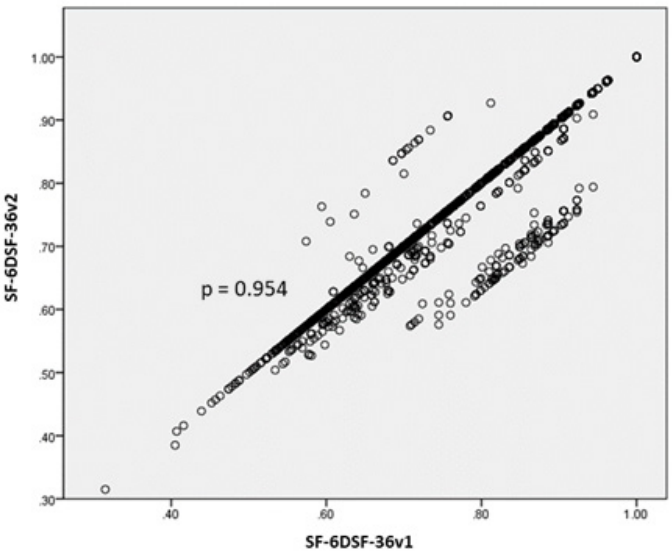
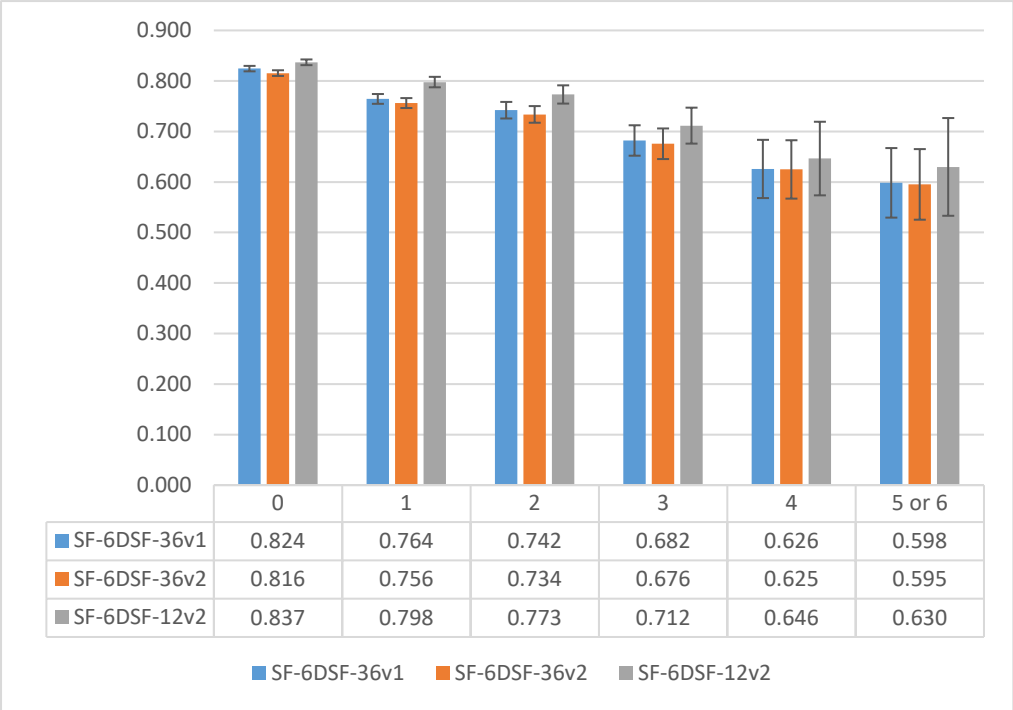



Figure 4: Mean SF-6D health utility score and its 95% confidence interval stratified by the number of chronic diseases




| | 0 | 1 | 2 | 3 | 4 | 5 or 6 |
|---|---|---|---|---|---|---|
| SF-6DSF-36v1 | 0.824 | 0.764 | 0.742 | 0.682 | 0.626 | 0.598 |
| SF-6DSF-36v2 | 0.816 | 0.756 | 0.734 | 0.676 | 0.625 | 0.595 |
| SF-6DSF-12v2 | 0.837 | 0.798 | 0.773 | 0.712 | 0.646 | 0.630 |

Table 1. Descriptive statistics of SF-6D scores derived by three Health Surveys by gender and age

| | n, % | $SF\text{-}6D_{SF\text{-}36v1}$ | | | $SF\text{-}6D_{SF\text{-}36v2}$ | | | $SF\text{-}6D_{SF\text{-}12v2}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | SD | 95% CI | Mean | SD | 95% CI | Mean | SD | 95% CI |
| Total | 2410 | 0.795 | 0.120 | (0.7899, 0.7996) | 0.786 | 0.123 | (0.7813, 0.7911) | 0.815 | 0.126 | (0.8097, 0.8198) |
| **Gender** | | | | | | | | | | |
| Male | 1258, 52.5% | 0.808 | 0.115 | (0.8011, 0.8144) | 0.799 | 0.118 | (0.7924, 0.8060) | 0.824 | 0.119 | (0.8173, 0.8311) |
| Female | 1152, 47.8% | 0.783 | 0.124 | (0.7760, 0.7897) | 0.774 | 0.126 | (0.7674, 0.7813) | 0.806 | 0.130 | (0.7989, 0.8133) |
| P-value | | P < 0.001* | | | P < 0.001* | | | P < 0.001* | | |
| **Age** | | | | | | | | | | |
| 18-20 | 241, 10.0% | 0.799 | 0.112 | (0.7848, 0.8132) | 0.786 | 0.116 | (0.7708, 0.8002) | 0.797 | 0.119 | (0.7814, 0.8116) |
| 21-30 | 439, 18.2% | 0.811 | 0.112 | (0.8008, 0.8218) | 0.806 | 0.114 | (0.7949, 0.8163) | 0.823 | 0.113 | (0.8120, 0.8332) |
| 31-40 | 564, 23.4% | 0.797 | 0.119 | (0.7870, 0.8067) | 0.789 | 0.122 | (0.7789, 0.7991) | 0.816 | 0.120 | (0.8063, 0.8261) |
| 41-50 | 357, 14.8% | 0.794 | 0.118 | (0.7818, 0.8063) | 0.787 | 0.121 | (0.7746, 0.7997) | 0.826 | 0.123 | (0.8132, 0.8388) |
| 51-60 | 241, 10.0% | 0.801 | 0.124 | (0.7854, 0.8169) | 0.790 | 0.127 | (0.7740, 0.8062) | 0.831 | 0.123 | (0.8154, 0.8467) |
| 61-70 | 266, 11.0% | 0.770 | 0.126 | (0.7551, 0.7855) | 0.765 | 0.128 | (0.7495, 0.7805) | 0.807 | 0.139 | (0.7901, 0.8238) |
| 71-80 | 162, 6.7% | 0.769 | 0.143 | (0.7467, 0.7911) | 0.765 | 0.144 | (0.7426, 0.7874) | 0.795 | 0.157 | (0.7709, 0.8196) |
| 81-90 | 38, 1.6% | 0.746 | 0.135 | (0.7013, 0.7898) | 0.742 | 0.132 | (0.6989, 0.7859) | 0.773 | 0.159 | (0.7207, 0.8252) |
| P-value | | P < 0.001† | | | P < 0.001† | | | P = 0.002† | | |

SD = standard deviation; CI = confidence interval

Note:

* Difference in SF-6D scores between males and females by independent t-test

† Difference in SF-6D scores between age group by analysis of variance

Table 2. Descriptive statistics of SF-6D scores derived from three Health Surveys by health condition

| | $SF\text{-}6D_{SF\text{-}36v1}$ | | | $SF\text{-}6D_{SF\text{-}36v2}$ | | | $SF\text{-}6D_{SF\text{-}12v2}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | 95% CI | Mean | SD | 95% CI | Mean | SD | 95% CI |
| Total (n=2410) | 0.795 | 0.120 | (0.7899, 0.7996) | 0.786 | 0.123 | (0.7813, 0.7911) | 0.815 | 0.126 | (0.8097, 0.8198) |
| Aged 60 or above (n=516) | 0.794 | 0.121 | (0.7890, 0.7989) | 0.786 | 0.124 | (0.7814, 0.7914) | 0.815 | 0.126 | (0.8100, 0.8203) |
| No chronic diseases (n=1493) | 0.824 | 0.106 | (0.8190, 0.8298) | 0.815 | 0.111 | (0.8098, 0.8211) | 0.837 | 0.111 | (0.8313, 0.8426) |
| Any chronic diseases (n=917) | 0.746 | 0.126 | (0.7383, 0.7546) | 0.739 | 0.127 | (0.7304, 0.7469) | 0.779 | 0.138 | (0.7696, 0.7876) |
| Hypertension (n=271) | 0.746 | 0.138 | (0.7299, 0.7629) | 0.739 | 0.140 | (0.7227, 0.7561) | 0.773 | 0.156 | (0.7546, 0.7919) |
| Stroke (n=21) | 0.710 | 0.164 | (0.6359, 0.7849) | 0.707 | 0.165 | (0.6314, 0.7819) | 0.737 | 0.178 | (0.6560, 0.8176) |
| Diabetes mellitus (n=110) | 0.736 | 0.139 | (0.7095, 0.7622) | 0.727 | 0.140 | (0.7001, 0.7531) | 0.759 | 0.160 | (0.7286, 0.7889) |
| Arthritis or other chronic rheumatism (n=473) | 0.731 | 0.128 | (0.7195, 0.7426) | 0.721 | 0.129 | (0.7091, 0.7324) | 0.765 | 0.144 | (0.7523, 0.7784) |
| Chronic lung disease (asthma, emphysema, chronic bronchitis or other chronic lung diseases) (n=128) | 0.719 | 0.132 | (0.6957, 0.7419) | 0.717 | 0.130 | (0.6937, 0.7393) | 0.745 | 0.140 | (0.7205, 0.7695) |
| Any heart disease (n=94) | 0.698 | 0.131 | (0.6716, 0.7253) | 0.692 | 0.127 | (0.6656, 0.7176) | 0.728 | 0.159 | (0.6955, 0.7606) |
| Mental illness (depression, anxiety, neurasthenia, psychosis or other mental illness) (n=94) | 0.674 | 0.112 | (0.6514, 0.6974) | 0.667 | 0.111 | (0.6446, 0.6901) | 0.715 | 0.135 | (0.6874, 0.7428) |
| Any other major disease (n=225) | 0.734 | 0.130 | (0.7169, 0.7512) | 0.730 | 0.132 | (0.7127, 0.7475) | 0.768 | 0.138 | (0.7495, 0.7857) |

SD = standard deviation; CI = confidence interval

Table 3. Proportion of respondents with no problems in each SF-6D dimension

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **SF-36v1** | | | | | | | | |
| | Male | | | | | | | | |
| Age group | All | 18-20 | 21-30 | 31-40 | 41-50 | 51-60 | 61-70 | 71-80 | 81-90 |
| Number of n | 1152 | 118 | 233 | 294 | 153 | 101 | 129 | 77 | 19 |
| Physical Functioning | 62.3% | 84.7% | 83.7% | 67.3% | 59.5% | 47.5% | 37.2% | 26.0% | 10.5% |
| Role Limitation | 68.7% | 61.0% | 62.7% | 66.7% | 73.9% | 82.2% | 69.8% | 68.8% | 84.2% |
| Social Functioning | 79.2% | 71.2% | 78.1% | 74.8% | 85.0% | 87.1% | 84.5% | 79.2% | 89.5% |
| Pain | 59.1% | 59.3% | 60.5% | 57.5% | 61.4% | 60.4% | 54.3% | 62.3% | 47.4% |
| Mental Health | 24.6% | 26.1% | 18.1% | 19.3% | 20.0% | 25.7% | 38.6% | 44.7% | 44.4% |
| Vitality | 9.6% | 7.4% | 7.1% | 8.0% | 9.5% | 14.1% | 12.6% | 12.9% | 20.0% |
| | Female | | | | | | | | |
| Age group | All | 18-20 | 21-30 | 31-40 | 41-50 | 51-60 | 61-70 | 71-80 | 81-90 |
| Number of n | 1258 | 123 | 206 | 270 | 204 | 140 | 137 | 85 | 19 |
| Physical Functioning | 46.7% | 80.5% | 72.8% | 51.5% | 39.2% | 30.0% | 19.0% | 12.9% | 5.3% |
| Role Limitation | 65.3% | 57.7% | 67.0% | 67.4% | 67.6% | 61.4% | 62.8% | 64.7% | 57.9% |
| Social Functioning | 78.5% | 67.5% | 82.5% | 74.4% | 82.4% | 82.1% | 80.3% | 77.6% | 78.9% |
| Pain | 50.6% | 55.3% | 63.6% | 54.1% | 49.5% | 45.7% | 36.5% | 32.9% | 31.6% |
| Mental Health | 21.5% | 13.4% | 19.7% | 17.7% | 17.5% | 29.2% | 29.0% | 31.8% | 38.9% |
| Vitality | 8.4% | 7.1% | 3.7% | 9.2% | 9.1% | 13.4% | 7.5% | 11.6% | 6.3% |
| | **SF-36v2** | | | | | | | | |
| | Male | | | | | | | | |
| Age group | All | 18-20 | 21-30 | 31-40 | 41-50 | 51-60 | 61-70 | 71-80 | 81-90 |
| Number of n | 1152 | 118 | 233 | 294 | 153 | 101 | 129 | 77 | 19 |
| Physical Functioning | 62.3% | 84.7% | 83.7% | 67.3% | 59.5% | 47.5% | 37.2% | 26.0% | 10.5% |
| Role Limitation | 63.5% | 53.4% | 59.2% | 61.2% | 69.9% | 75.2% | 67.4% | 66.2% | 73.7% |
| Social Functioning | 79.2% | 71.2% | 78.1% | 74.8% | 85.0% | 87.1% | 84.5% | 79.2% | 89.5% |
| Pain | 59.1% | 59.3% | 60.5% | 57.5% | 61.4% | 60.4% | 54.3% | 62.3% | 47.4% |
| Mental Health | 24.0% | 24.6% | 18.0% | 18.7% | 19.6% | 25.7% | 38.0% | 44.2% | 42.1% |
| Vitality | 7.5% | 5.9% | 5.6% | 6.1% | 7.2% | 9.9% | 10.1% | 10.4% | 15.8% |
| | Female | | | | | | | | |
| Age group | All | 18-20 | 21-30 | 31-40 | 41-50 | 51-60 | 61-70 | 71-80 | 81-90 |
| Number of n | 1258 | 123 | 206 | 270 | 204 | 140 | 137 | 85 | 19 |
| Physical Functioning | 46.7% | 80.5% | 72.8% | 51.5% | 39.2% | 30.0% | 19.0% | 12.9% | 5.3% |
| Role Limitation | 59.9% | 50.4% | 64.1% | 63.3% | 62.7% | 57.9% | 58.4% | 61.2% | 57.9% |
| Social Functioning | 78.5% | 67.5% | 82.5% | 74.4% | 82.4% | 82.1% | 80.3% | 77.6% | 78.9% |
| Pain | 50.6% | 55.3% | 63.6% | 54.1% | 49.5% | 45.7% | 36.5% | 32.9% | 31.6% |
| Mental Health | 21.0% | 13.0% | 18.9% | 17.4% | 17.2% | 28.6% | 27.7% | 31.8% | 36.8% |
| Vitality | 6.8% | 5.7% | 2.9% | 7.0% | 7.4% | 12.1% | 5.8% | 9.4% | 5.3% |
| | **SF-12v2** | | | | | | | | |
| | Male | | | | | | | | |

| Age group | All | 18-20 | 21-30 | 31-40 | 41-50 | 51-60 | 61-70 | 71-80 | 81-90 |
|---|---|---|---|---|---|---|---|---|---|
| Number of n | 1152 | 118 | 233 | 294 | 153 | 101 | 129 | 77 | 19 |
| Physical Functioning | 91.3% | 99.2% | 97.9% | 96.6% | 92.2% | 92.1% | 79.1% | 68.8% | 42.1% |
| Role Limitation | 63.5% | 53.4% | 59.2% | 61.2% | 69.9% | 75.2% | 67.4% | 66.2% | 73.7% |
| Social Functioning | 79.2% | 71.2% | 78.1% | 74.8% | 85.0% | 87.1% | 84.5% | 79.2% | 89.5% |
| Pain | 83.1% | 86.4% | 85.8% | 79.9% | 85.0% | 82.2% | 82.2% | 77.9% | 73.7% |
| Mental Health | 32.3% | 33.9% | 26.2% | 26.9% | 32.0% | 33.7% | 44.2% | 49.4% | 47.4% |
| Vitality | 7.5% | 5.9% | 5.6% | 6.1% | 7.2% | 9.9% | 10.1% | 10.4% | 15.8% |
| | | | | | Female | | | | |
| Age group | All | 18-20 | 21-30 | 31-40 | 41-50 | 51-60 | 61-70 | 71-80 | 81-90 |
| Number of n | 1258 | 123 | 206 | 270 | 204 | 140 | 137 | 85 | 19 |
| Physical Functioning | 85.9% | 96.7% | 97.1% | 93.3% | 87.7% | 84.3% | 69.3% | 50.6% | 36.8% |
| Role Limitation | 59.9% | 50.4% | 64.1% | 63.3% | 62.7% | 57.9% | 58.4% | 61.2% | 57.9% |
| Social Functioning | 78.5% | 67.5% | 82.5% | 74.4% | 82.4% | 82.1% | 80.3% | 77.6% | 78.9% |
| Pain | 76.5% | 78.0% | 85.4% | 79.3% | 79.4% | 69.3% | 65.7% | 65.9% | 63.2% |
| Mental Health | 28.5% | 19.5% | 25.7% | 27.4% | 25.5% | 36.4% | 33.6% | 37.6% | 42.1% |
| Vitality | 6.8% | 5.7% | 2.9% | 7.0% | 7.4% | 12.1% | 5.8% | 9.4% | 5.3% |

Table 4. Response distribution of SF-6D response levels mapped from three Health Surveys

| | **SF-36v1** | | | | | |
|---|---|---|---|---|---|---|
| Level | Physical Functioning | Role Limitation | Social Functioning | Pain | Mental Health | Vitality |
| 1 (No problems) | 54.1% | 66.9% | 78.8% | 54.7% | 22.4% | 7.1% |
| 2 | 34.2% | 7.8% | 13.8% | 24.9% | 33.4% | 27.0% |
| 3 | 8.3% | 17.0% | 5.1% | 11.0% | 35.5% | 44.6% |
| 4 | 2.4% | 8.3% | 1.1% | 5.7% | 7.3% | 13.0% |
| 5 | 0.7% | N/A | 1.1% | 2.4% | 1.5% | 8.3% |
| 6 | 0.2% | N/A | N/A | 1.3% | N/A | N/A |
| | **SF-36v2** | | | | | |
| Level | Physical Functioning | Role Limitation | Social Functioning | Pain | Mental Health | Vitality |
| 1 (No problems) | 54.1% | 61.6% | 78.8% | 54.7% | 22.4% | 7.1% |
| 2 | 34.2% | 8.8% | 13.8% | 24.9% | 33.4% | 26.9% |
| 3 | 8.3% | 17.4% | 5.1% | 11.0% | 35.2% | 44.7% |
| 4 | 2.4% | 12.3% | 1.1% | 5.7% | 7.5% | 13.0% |
| 5 | 0.7% | N/A | 1.1% | 2.4% | 1.5% | 8.3% |
| 6 | 0.2% | N/A | N/A | 1.3% | N/A | N/A |
| | **SF-12v2** | | | | | |
| Level | Physical Functioning | Role Limitation | Social Functioning | Pain | Mental Health | Vitality |
| 1 (No problems) | 88.5% | 61.6% | 78.8% | 79.6% | 30.3% | 7.1% |
| 2 | 8.6% | 8.8% | 13.8% | 11.0% | 39.3% | 26.9% |
| 3 | 2.9% | 17.4% | 5.1% | 5.7% | 26.1% | 44.7% |
| 4 | N/A | 12.3% | 1.1% | 2.4% | 3.6% | 13.0% |
| 5 | N/A | N/A | 1.1% | 1.3% | 0.6% | 8.3% |
| 6 | N/A | N/A | N/A | N/A | N/A | N/A |

N/A = Not applicable

# SF-12v2 Conversion to SF-6D Scoring Algorithm

| **Physical Functioning (PF) Item response** | **(Choose 1 only)** | ***SF-6D PF Value*** |
|---|---|---|
| SF-6D PF 1 (SF-12v2 Q2a = 3) | | *0* |
| SF-6D PF 2 (SF-12v2 Q2a = 2) | | *-0.045* |
| SF-6D PF 3 (SF-12v2 Q2a = 1)* | | *-0.129* |
| **Role Limitation (RL) Item response** | **(Choose 1 only)** | ***SF-6D RL Value*** |
| SF-6D RL1 (SF-12v2 Q3b = 5 & Q4a = 5) | | *0* |
| SF-6D RL2 (SF-12v2 Q3b < 5) | | *-0.033* |
| SF-6D RL3 (SF-12v2 Q4a < 5)* | | *-0.036* |
| SF-6D RL4 (SF-12v2 Q3b < 5 & Q4a < 5)* | | *-0.055* |
| **Social Functioning (SF) Item response** | **(Choose 1 only)** | ***SF-6D SF Value*** |
| SF-6D SF1 (SF-12v2 Q7 = 5) | | *0* |
| SF-6D SF2 (SF-12v2 Q7 = 4) | | *-0.041* |
| SF-6D SF3 (SF-12v2 Q7 = 3) | | *-0.048* |
| SF-6D SF4 (SF-12v2v2 Q7 = 2)* | | *-0.050* |
| SF-6D SF5 (SF-12v2v2 Q7 = 1)* | | *-0.070* |
| **Pain (Pain) Item response** | **(Choose 1 only)** | ***SF-6D Pain Value*** |
| SF-6D BP1 (SF-12v2 Q5 = 1) | | *0* |
| SF-6D BP2 (SF-12v2 Q5 = 2) | | *0* |
| SF-6D BP3 (SF-12v2 Q5 = 3) | | *-0.038* |
| SF-6D BP4 (SF-12v2 Q5 = 4)* | | *-0.038* |
| SF-6D BP5 (SF-12v2 Q5 = 5)* | | *-0.082* |
| **Mental Health (MH) Item response** | **(Choose 1 only)** | ***SF-6D MH Value*** |
| SF-6D MH1 (SF-12v2 Q6c = 5) | | *0* |
| SF-6D MH2 (SF-12v2 Q6c = 4) | | *-0.044* |
| SF-6D MH3 (SF-12v2 Q6c = 3) | | *-0.068* |
| SF-6D MH4 (SF-12v2 Q6c = 2)* | | *-0.087* |
| SF-6D MH5 (SF-12v2 Q6c = 1)* | | *-0.091* |
| **Vitality (VT) Item response** | **(Choose 1 only)** | ***SF-6D VT Value*** |
| SF-6D VT1 (SF-12v2 Q6b = 1) | | *0* |
| SF-6D VT2 (SF-12v2 Q6b = 2) | | *-0.043* |
| SF-6D VT3 (SF-12v2 Q6b = 3) | | *-0.063* |
| SF-6D VT4 (SF-12v2 Q6b = 4) | | *-0.073* |
| SF-6D VT5 (SF-12v2 Q6b = 5)* | | -0.076 |
| **** Add MOST Adjustment Value*** **(applicatble if any dimension value marked by * is endorsed)** | | *-0.139* |

***SF-6D Index score ( 1= perfect health, 0=death)***

= ***1 + PF value + RL value + BP value + SF value + MH value +VT value + MOST value (if applicable) =***

# SF-36v1 Conversion to SF-6D Scoring Algorithm

| **Physical Functioning (PF) Item response** | **(Choose 1 only)** | ***SF-6D PF Value*** |
|---|---|---|
| SF-6D PF 1 (SF-36v2 Q3a=3) | | *0* |
| SF-6D PF 2 (SF-36v2 Q3a=2/1) | | *-0.050* |
| SF-6D PF 3 (SF-36v2 Q3b=2) | | *-0.056* |
| SF-6D PF 4 (SF-36v2 Q3b=1)* | | *-0.092* |
| SF-6D PF 5 (SF-36v2 Q3j=2)* | | *-0.103* |
| SF-6D PF 6 (SF-36v2 Q3j=1)* | | *-0.178* |
| **Role Limitation (RL) Item response** | **(Choose 1 only)** | ***SF-6D RL Value*** |
| SF-6D RL1 (SF-36v2 Q4c=2 & Q5b = 2) | | *0* |
| SF-6D RL2 (SF-36v2 Q4c=1 & Q5b = 2) | | *-0.035* |
| SF-6D RL3 (SF-36v2 Q5b=1 & Q4c = 2)* | | *-0.035* |
| SF-6D RL4 (SF-36v2 Q4c=1 & Q5b = 1)* | | *-0.054* |
| **Social Functioning (SF) Item response** | **(Choose 1 only)** | ***SF-6D SF Value*** |
| SF-6D SF1 (SF-36v2 Q10=5) | | *0* |
| SF-6D SF2 (SF-36v2 Q10=4) | | *-0.039* |
| SF-6D SF3 (SF-36v2 Q10=3) | | *-0.050* |
| SF-6D SF4 (SF-36v2 Q10=2)* | | *-0.050* |
| SF-6D SF5 (SF-36v2 Q10=1)* | | *-0.073* |
| **Pain (Pain) Item response** | **(Choose 1 only)** | ***SF-6D Pain Value*** |
| SF-6D BP1 (SF-36v2 Q7=1) | | *0* |
| SF-6D BP2 (SF-36v2 Q7≥2 & Q8=1) | | *-0.037* |
| SF-6D BP3 (SF-36v2 Q8=2) | | *-0.037* |
| SF-6D BP4 (SF-36v2 Q8=3) | | *-0.052* |
| SF-6D BP5 (SF-36v2 Q8=4)* | | *-0.060* |
| SF-6D BP6 (SF-36v2 Q8=5)* | | *-0.100* |
| **Mental Health (MH) Item response †** | **(Choose 1 only)** | ***SF-6D MH Value*** |
| SF-6D MH1 (SF-36v2 Q9b=6 & Q9f=6) | | *0* |
| SF-6D MH2 (SF-36v2 Q9b=5 or Q9f=5) | | *-0.038* |
| SF-6D MH3 (SF-36v2 Q9b=4 or Q9f=4) | | *-0.058* |
| SF-6D MH4 (SF-36v2 Q9b=2 or Q9f=2)* | | *-0.088* |
| SF-6D MH5 (SF-36v2 Q9b=1 or Q9f=1)* | | *-0.088* |
| **Vitality (VT) Item response †** | **(Choose 1 only)** | ***SF-6D VT Value*** |
| SF-6D VT1 (SF-36v2 Q9e=1) | | *0* |
| SF-6D VT2 (SF-36v2 Q9e=2) | | *-0.039* |
| SF-6D VT3 (SF-36v2 Q9e=4) | | *-0.056* |
| SF-6D VT4 (SF-36v2 Q9e=5)* | | *-0.063* |
| SF-6D VT5 (SF-36v2 Q9e=6)* | | *-0.077* |
| *** Add MOST Adjustment Value** ***(applicable if any dimension value marked by * is endorsed)*** | | *-0.115* |
| **† The response value "3" for question 9b, 9e and 9e items was recoded randomly to the adjoining values of "2" or "4".** | | |

***SF-6D Index score ( 1= perfect health, 0=death)***

= ***1 + PF value + RL value + BP value + SF value + MH value +VT value + MOST value (if applicable) =***

# SF-36v2 Conversion to SF-6D Scoring Algorithm

| **Physical Functioning (PF) Item response** | **(Choose 1 only)** | ***SF-6D PF Value*** |
|---|---|---|
| SF-6D PF 1 (SF-36v2 Q3a=3) | | *0* |
| SF-6D PF 2 (SF-36v2 Q3a=2/1) | | *-0.050* |
| SF-6D PF 3 (SF-36v2 Q3b=2) | | *-0.056* |
| SF-6D PF 4 (SF-36v2 Q3b=1)* | | *-0.092* |
| SF-6D PF 5 (SF-36v2 Q3j=2)* | | *-0.103* |
| SF-6D PF 6 (SF-36v2 Q3j=1)* | | *-0.178* |
| **Role Limitation (RL) Item response** | **(Choose 1 only)** | ***SF-6D RL Value*** |
| SF-6D RL1 (SF-36v2 Q4c=5 & Q5b = 5) | | *0* |
| SF-6D RL2 (SF-36v2 Q4c<5 & Q5b = 5) | | *-0.035* |
| SF-6D RL3 (SF-36v2 Q5b<5 & Q4c = 5)* | | *-0.035* |
| SF-6D RL4 (SF-36v2 Q4c<5 & Q5b<5)* | | *-0.054* |
| **Social Functioning (SF) Item response** | **(Choose 1 only)** | ***SF-6D SF Value*** |
| SF-6D SF1 (SF-36v2 Q10=5) | | *0* |
| SF-6D SF2 (SF-36v2 Q10=4) | | *-0.039* |
| SF-6D SF3 (SF-36v2 Q10=3) | | *-0.050* |
| SF-6D SF4 (SF-36v2 Q10=2)* | | *-0.050* |
| SF-6D SF5 (SF-36v2 Q10=1)* | | *-0.073* |
| **Pain (Pain) Item response** | **(Choose 1 only)** | ***SF-6D Pain Value*** |
| SF-6D BP1 (SF-36v2 Q7=1) | | *0* |
| SF-6D BP2 (SF-36v2 Q7≥2 & Q8=1) | | *-0.037* |
| SF-6D BP3 (SF-36v2 Q8=2) | | *-0.037* |
| SF-6D BP4 (SF-36v2 Q8=3) | | *-0.052* |
| SF-6D BP5 (SF-36v2 Q8=4)* | | *-0.060* |
| SF-6D BP6 (SF-36v2 Q8=5)* | | *-0.100* |
| **Mental Health (MH) Item response** | **(Choose 1 only)** | ***SF-6D MH Value*** |
| SF-6D MH1 (SF-36v2 Q9b=5 & Q9f=5) | | *0* |
| SF-6D MH2 (SF-36v2 Q9b=4 or Q9f=4) | | *-0.038* |
| SF-6D MH3 (SF-36v2 Q9b=3 or Q9f=3) | | *-0.058* |
| SF-6D MH4 (SF-36v2 Q9b=2 or Q9f=2)* | | *-0.088* |
| SF-6D MH5 (SF-36v2 Q9b=1 or Q9f=1)* | | *-0.088* |
| **Vitality (VT) Item response** | **(Choose 1 only)** | ***SF-6D VT Value*** |
| SF-6D VT1 (SF-36v2 Q9e=1) | | *0* |
| SF-6D VT2 (SF-36v2 Q9e=2) | | *-0.039* |
| SF-6D VT3 (SF-36v2 Q9e=3) | | *-0.056* |
| SF-6D VT4 (SF-36v2 Q9e=4)* | | *-0.063* |
| SF-6D VT5 (SF-36v2 Q9e=5)* | | *-0.077* |
| *** Add MOST Adjustment Value** ***(applicable if any dimension value marked by * is endorsed)*** | | *-0.115* |

***SF-6D Index score ( 1= perfect health, 0=death)***

= ***1 + PF value + RL value + BP value + SF value + MH value +VT value + MOST value (if applicable) =***