

國人生活品質評量(II)： SF-36 台灣版的常模與效度檢測

曾旭民¹ 盧瑞芬^{1,*} 蔡益堅²

Hsu-Min Tseng¹, Jui-Fen Rachel Lu^{1,*}, Yih-Jian Tsai²

¹ 長庚大學醫務管理學系健康資料庫實驗室，桃園縣龜山鄉文化一路259號

Laboratory for Health Databank, Department of Health Care Management, Chang Gung University, 259 Wen-Hwa 1st Road, Kwei-Shan Tao-Yuan, Taiwan, R.O.C.

² 行政院衛生署國民健康局人口與健康調查研究中心

Center for Population and Health Survey Research, Bureau of Health Promotion, Department of Health, Taichung, Taiwan, R.O.C.

*通訊作者Correspondence author. E-mail: rachel@mail.cgu.edu.tw

目標：建立SF-36台灣版的常模，及進行初步的效標效度檢定，同時亦與美國常模作一比較。

方法：本研究以2001年國民健康調查12歲(含)以上之受訪者為分析樣本(N = 18,142)，其中共有17,515人進行SF-36的常模及效標效度檢定。**結果：**本研究發現SF-36台灣版與世界各國的研究結果相同，SF-36台灣版的常模資料顯示，男性在各量表上的平均值明顯高於女性。分析年齡與常模分數之間的關係顯示，屬於身體生理健康領域的量表分數隨著年齡增加而遞減；但屬於心理健康領域下的量表則較不受年齡影響。以年齡作為效標，比較25-34歲與55-64歲兩組年齡層在SF-36的表現，其測試結果亦符合原始版本的假設。**結論：**本研究所顯示的常模資料立基於具有全國代表性的樣本，其結果可以作為日後使用SF-36台灣版的比較依據。在效度的研究上，則尚需更多SF-36台灣版研究者的努力，以增進SF-36台灣版的解釋力。(台灣衛誌 2003；22(6)：512-518)

關鍵詞：健康生活品質，SF-36台灣版，常模，效度

Assessment of health-related quality of life in Taiwan (II): norming and validation of SF-36 Taiwan version

Objectives: This study aims to establish the norm of the SF-36 Taiwan version and to test criteria-validity underlying the scale construction. A comparison in norms was also made between US and Taiwan. **Methods:** The study sample was composed of 18,142 subjects aged 12 and above, collected from the 2001 Health Interview Survey. Norming and validation of SF-36 Taiwan version were performed on a valid sample of 17,515 subjects. **Results:** Our study results were consistent with those found in other language versions. Male subjects tend to score significantly higher than female subject on all scales. As age increases, the scales in the physical health dimension tend to perform worse, however, the age effect on the scales in the mental health dimension is less pronounced. Using age as a criteria, preliminary results from the criteria validity testing performed on younger (25-34y/o) and older (55-64y/o) adults found similar patterns to the US version. **Conclusion:** The paper presents the norm of the SF-36 Taiwan version which can serve as a valuable reference for future comparisons. However, more research in validation will warrant and enhance the explanatory power of the SF-36 Taiwan version. (Taiwan J Public Health. 2003; 22(6):512-518)

Keywords: health-related quality of life, SF-36 Taiwan version, norm, validity

前 言

雖然聯合國世界衛生組織在1946年成立的憲章中定義「健康」為生理、心理與社會上的安寧美好狀態，而不僅是免於疾病或虛弱。不過在傳統上，醫療的主要目的還是在於讓病患免於病痛之苦，因此過去在醫學治療上的結果研究(outcome study)較重視病人接受治療的前後客觀生理指標之改變。在晚近20年來，由於醫學的進步，使得死亡率降低及慢性疾病人口數增加，同時醫療人性化日益受到社會與醫療資源控管單位的重視，因此在醫療結果研究所選用的指標方面開始有顯著的變化。此變化主要是針對病患的健康狀態評估，除客觀的臨床生理指標外，同時更重視病患主觀上對生理、心理及社會三方面與健康有關的生活品質(Health-Related Quality of Life, HR-QoL)評估。

因應此發展趨勢，從80年代以來西方學者開始發展各種與健康相關的生活品質心理測量工具，而這些量表也同時廣泛應用於如癌症、心血管疾病及精神科等各種不同疾病的病人[1]。HR-QoL依其適用的範圍主要可以區分成兩種：(1)一般型(generic)：主要適用在一般族群的調查，由於其通用性，也可以應用來測量各類疾病病患的健康狀態；(2)特定疾病型(disease-specific)：主要考量特定的疾病病患(如癌症或中風病人)，其身心功能的特殊性而加以設計。這些數量眾多的HR-QoL測驗工具主要都是從英語系國家發展而成，同時大部分有關心理計量上的特質如信度與效度的研究，亦都來自原始發展國家。直到最近10年，由於政府醫療管理單位的重視，同時藥廠跨國研究在進行藥物人體實驗時，皆須具有測量病人健康相關之生活品質的跨文化計量工具等因素，一般型HR-QoL逐漸有跨國與跨文化的研究[2,3]。

在一般型HR-QoL測量中，SF-36健康狀態調查表是最常被使用，且發行多國語文版的一種。SF-36是John Ware, Jr及其同儕於80年代末期在美國進行Medical Outcome Study

中所發展出來的多面向健康生活品質評估問卷，共分成8個量表：身體生理功能(physical functioning, PF)，因身體生理問題角色受限(role limitation due to physical health problems, RP)，身體疼痛(bodily pain, BP)，一般健康狀況(general health, GH)，活力狀況(vitality, VT)，社會功能(social functioning, SF)，因情緒問題角色受限(role limitation due to emotional problems, RE)，心理健康(mental health, MH)，多項研究皆顯示SF-36是一份具有良好信效度的評估工具[4]。為便利跨國文化比較研究，John Ware及其研究團隊在1991年進行國際生活品質評估(International Quality of Life Assessment, IQOLA)研究計畫，其研究結果顯示不同國家的SF-36版本具有良好的計量特質[5,6]。IQOLA研究計畫中設定的準則(protocol)在步驟上分為翻譯(translation)、對量表假設的檢測(tests of scaling assumptions)、實證研究建立效度(validation)、及常模(norm)的建立[7]。

中文化的SF-36在不同國家或區域有不同的版本，包含美國及香港地區的中文版本[8,9]。SF-36臺灣版本則由國內學者盧瑞芬依據IQOLA研究計畫準則於1996年發展並授權學術使用，發展過程詳見另篇論文[10]。雖然SF-36台灣版具有表面效度(face validity)，在發展後經過學者相繼應用，亦顯示該量表具有良好信度與效度[11,12]，但是這些研究所施測的對象屬特定病人族群或是樣本不具有台灣母體代表性，因此依據IQOLA研究計畫準則，SF-36台灣版尚無法與其他國家版本進行跨文化的比較研究。

任何量表心理計量的效度檢驗對該量表而言都是相當重要，因為效度研究可以增進對量表所得分數或分數改變的解釋力[13]。而常模的建立對預期應用於廣泛族群的測驗工具更形重要，因為常模資料可作為比較基準，有助於研究族群在健康狀態表現之比較，進而增進其研究上的解釋力。本篇論文的主要目的即為以2001年的國民健康調查(國健調查)資料，建立SF-36台灣版的常模，同時亦藉由不同年齡與性別之間的比較，建立

投稿日期：92年1月13日

接受日期：92年3月26日

此版本的初步效度指標。

材料與方法

資料與分析

關於國健調查的抽樣、樣本分配、SF-36施測與及本文中所選用SF-36有效資料選擇的詳細描述，請參見本期另篇論文[10]。其中共有17,276筆資料有效資料來進行常模與效度方面的相關分析，其中男性有8,671筆(49.5%)、女性為8,844筆。

在常模的建立方面，原則上依據美國常模呈現的方式以便利於相互比較。首先將各量表的原始總分依據轉換原則，轉換成從0到100的總分，其轉換的公式為：

轉換分數 = [(量表原始分數 - 可能最低分數) / (量表可能的分數間距)] × 100。而在年齡區分上則以10歲為一個年齡區段，對於小於25歲的樣本則區分為12-17歲以及18-24歲。其次我們進行SF-36台灣版常模與美國版常模的比較，以瞭解兩者之間的異同。在效度分析方面，則進行效標效度(criteria validity)檢驗。以年齡作為效標，根據原量表的假設，老年人與年輕人的主要差異應該在於身體生理健康領域(physical health dimension)有關的量表(如PF, RP, 與BP)，而在心理健康領域(mental health dimension)的量表(如MH, RE, 與SF)，則較不受年齡的影響。在效標效度分析上，根據美國版使用手冊中在身體生理健康領域方面，55歲以上的年齡群對比於較年輕的年齡層開始有較大的降幅，因此我們選取年輕成人(25到34歲間，共3,318人)與較老成人(55到64歲間，共1,663人)兩個年齡層加以比較來檢驗效標效度。

結 果

SF-36台灣版依據年齡與性別區分的常模結果如表一所示。在性別的比較上，男性在所有量表的平均分數都顯著高於女性($p < .0001$)，而且此種差異在老年人口(> 65歲)比青(少)年人口更為顯著。此外，隨著年齡的

增加，PF、RP、BP、GH與VT量表上的平均分數明顯隨之遞減；量表得分與年齡之間的關係在量表SF、RE與MH則較不明顯。

圖一顯示台灣與美國常模之間的比較結果。如圖所示，相較於美國常模，台灣版在量表PF、RP、BP、VT、與SF之平均值較高，而在量表GH、RE、與MH則較低；其中平均值差異較大者為量表PF、BP、與VT。由於兩個常模的人數眾多，因此這些差異在檢定上均顯示出有顯著的差異(t 值介於2.6到16.7之間， $p < .0001$)。

在效標效度檢驗部分，則主要是以年齡為效標以檢驗年齡對SF-36中身體生理健康與心理健康領域的影響。如圖二所示，分析結果顯示SF-36台灣版符合原版假設，亦即屬於身體生理健康領域的量表，如PF、RP、BP、GH與VT最容易受到年齡的影響(t 值介於10.6到28.0，所有 p 值都 < 0.001)，而屬於心理健康領域量表，如SF($t = 7.0$ ， $p < 0.001$)，RE($t = 2.0$ ， $p < 0.05$)，與MH($t = 1.8$ ， $p = 0.07$)則較不受年齡的影響，其中MH量表在兩個年齡層的差異不具有統計上的顯著性。

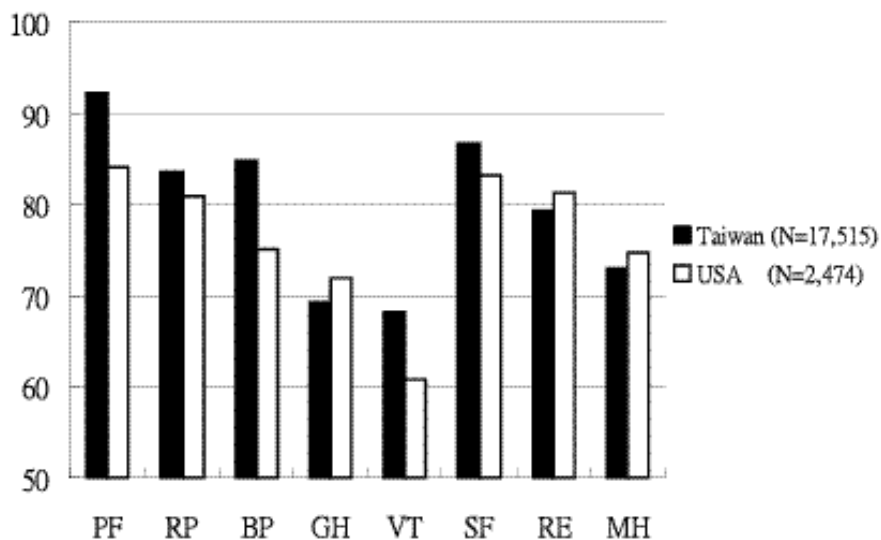
討 論

根據IQOLA研究計畫準則，常模研究所顯示的描述型數據具有非常重要的解釋價值[2]。一個具有全國代表性的常模結果，將有助於日後研究者或臨床工作者在解釋個人或特定患者施測之SF-36分數所代表的意義。本文以具全國代表性的2001年國健調查資料進行SF-36台灣版的常模與初步效度分析，分析結果顯示，SF-36台灣版常模分布趨勢與美國原版一致，同時在效標效度方面，亦符合美國原版SF-36的原始設計概念。

相較於其它SF-36中文版的研究[8,9,11,14,15]，本文的分析是目前所有SF-36中文版研究中，施測人數最多並且最具有全國代表性的研究。其它中文版本的研究，有的施測人數過少，例如香港的常模研究[9]僅施測236人，而其結果在各量表與其他國家版本比較也具有較大的差異；有的中文版本則使用社區抽樣或施測於特定的族群，如美國中文

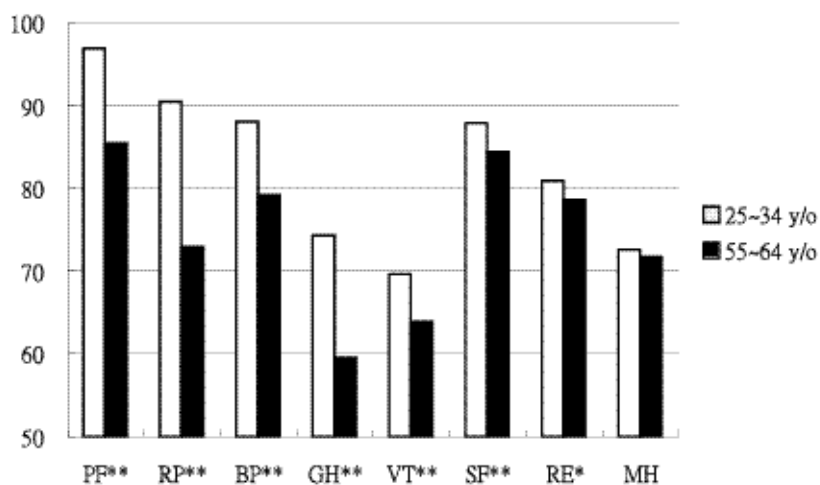
表一 SF-36台灣版常模資料(依性別與年齡分佈)

整體樣本	17515	92.24(16.16)	83.65(33.27)	84.84(19.42)	69.29(21.27)	68.27(18.66)	86.81(17.05)	79.4 (36.07)	73.01(16.55)
男	8671	93.99(14.62)	86.44(30.79)	87.59(18.06)	71.55(20.28)	70.95(17.90)	87.86(16.56)	81.24(34.65)	75.08(15.95)
女	8844	90.52(17.36)	80.91(35.31)	82.14(20.32)	67.08(21.99)	65.64(19.02)	85.78(17.46)	77.59(37.33)	70.99(16.88)
性別與年齡區分									
12~17歲	1827	98.20 (7.06)	91.98(22.22)	91.24(14.11)	77.56(17.63)	74.27(16.56)	88.64(14.83)	79.51(34.52)	74.02(15.01)
男	928	98.51 (6.01)	93.10(20.66)	92.83(12.57)	79.02(16.49)	75.40(15.82)	89.49(14.49)	81.03(33.63)	75.26(14.08)
女	899	97.86 (6.78)	90.82(25.08)	89.61(15.31)	76.04(17.65)	73.11(16.54)	87.75(15.90)	77.94(34.82)	72.74(15.09)
18~24歲	2499	97.84 (6.64)	89.50(26.48)	88.82(16.28)	75.11(18.79)	71.20(17.20)	87.75(15.41)	78.18(35.70)	73.06(15.73)
男	1266	98.26 (8.13)	90.78(22.78)	90.68(15.74)	77.35(17.20)	73.14(15.70)	88.29(15.17)	79.54(34.11)	74.44(15.05)
女	1233	97.40 (9.24)	88.18(26.92)	86.92(17.83)	72.81(19.02)	69.22(17.46)	87.19(15.49)	76.78(33.40)	71.65(16.22)
25~34歲	3318	96.83 (9.08)	90.48(25.33)	88.04(16.79)	74.33(18.10)	69.59(16.70)	87.84(15.08)	80.87(34.56)	72.56(15.29)
男	1623	97.94(11.59)	92.50(28.76)	90.33(17.82)	75.90(20.10)	72.36(18.02)	88.38(14.82)	81.52(31.28)	74.02(16.49)
女	1695	95.77(18.53)	88.55(39.28)	85.84(21.06)	72.83(22.09)	66.94(20.12)	87.32(18.04)	80.26(35.44)	71.16(17.36)
35~44歲	3500	95.50(10.57)	87.61(29.29)	85.58(18.22)	70.94(19.91)	68.81(18.24)	88.27(15.48)	82.27(33.96)	73.58(16.40)
男	1719	96.91(21.79)	89.85(42.27)	87.78(21.99)	73.16(22.46)	71.92(19.04)	88.69(19.98)	83.08(38.44)	75.59(17.35)
女	1781	94.14(29.04)	85.44(45.97)	83.45(23.97)	68.79(21.51)	65.81(21.98)	87.86(25.38)	81.49(44.45)	71.55(17.63)
45~54歲	2866	92.43(13.65)	83.85(32.74)	83.80(19.78)	66.60(21.26)	68.19(18.96)	88.18(16.16)	82.36(34.00)	74.09(17.57)
男	1391	94.74 (8.00)	87.85(23.67)	86.87(15.37)	69.36(18.61)	71.42(17.22)	89.69(15.13)	85.00(35.37)	76.66(15.82)
女	1475	90.25 (6.47)	80.08(27.80)	80.89(17.02)	64.01(19.63)	65.13(17.65)	86.75(14.87)	79.88(36.54)	71.66(16.24)
55~64歲	1663	85.45(19.16)	72.85(40.52)	79.21(21.76)	59.59(22.56)	63.89(20.31)	84.35(19.20)	78.67(37.98)	71.67(18.05)
男	788	88.27 (9.79)	75.76(27.43)	82.67(17.47)	62.04(18.80)	66.59(17.19)	86.25(14.97)	81.85(34.98)	74.39(15.38)
女	875	82.91(11.56)	70.23(31.26)	76.10(18.34)	57.38(20.51)	61.46(18.49)	82.64(15.46)	75.81(34.49)	69.21(16.35)
65~74歲	1269	75.04(24.20)	62.77(44.80)	74.64(23.65)	55.06(23.12)	60.76(19.89)	81.61(21.58)	72.97(41.17)	71.77(18.14)
男	648	80.82(15.02)	70.76(35.68)	80.52(21.07)	58.94(21.98)	64.92(19.31)	84.92(17.20)	77.83(36.20)	75.82(18.20)
女	621	69.02(19.36)	54.43(41.45)	68.50(21.93)	51.02(22.77)	56.42(20.18)	78.16(20.04)	67.90(39.93)	67.54(18.32)
75歲以上	573	59.15(29.30)	44.46(45.86)	68.08(25.42)	51.57(21.69)	55.25(21.35)	73.82(26.13)	59.80(46.47)	70.34(18.39)
男	308	64.74(25.13)	51.54(45.86)	74.21(23.78)	54.19(23.13)	57.61(19.83)	75.77(22.65)	65.80(43.30)	72.98(18.00)
女	265	52.66(28.29)	36.23(44.42)	60.96(25.24)	48.53(21.54)	52.51(20.29)	71.56(26.84)	52.83(47.86)	67.27(18.82)



圖一 SF-36各量表的台灣常模與美國常模比較

資料來源：美國常模資料來自Ware J, et al. [4]；台灣常模係作者自行計算而得。



圖二 SF-36台灣版各量表測量結果依年齡層比較

* $p < .05$, ** $p < .0001$



版[15]，因此基本上並不具有當地區域的代表性。有鑑於此，本研究並不特別與各地SF-36中文版相互比較，而僅進行與美國常模之間的比較。

與美國常模比較結果顯示，SF-36台灣版各量表的平均值在身體生理健康領域的5個量表(PF、RP、BP)，台灣的平均分數高於美國；心理健康領域的3個量表中(SF、RE、MH)，則除了SF，台灣的平均值均低於美國；此外兩個屬於一般健康方面的量表(VT與GH)則表現互異，在VT方面台灣高於美國，而GH則相反。需要說明的是這些差異原則上僅是描述統計上的敘述，並未進行統計上顯著性的檢測。正如SF-36的國際比較一般[2]，必然會發現平均值之間的差異，而導致差異的因素，除可能反映出各個國家實際健康狀態差異外，可能尚包括受到問卷施測方法及不同文化對健康的知覺方式不同等因素的影響。

從美國與其他國家的比較研究[16-18]中得知，在SF-36的各量表得分上，男性通常顯著高於女性的平均值，性別差異在SF-36台灣版上也得到相同的結果。不過與美國常模不同的是，在美國的常模中，性別差異在身體生理健康領域相關的量表比心理健康領域相關的量表更顯著，但在台灣常模中，性別差異則沒有如此顯著的差異。導致此種性別差異的原因，除可能顯示男女性別在真實的健康生活品質的不同外，從心理學的觀點而言，也許更有可能的是起因於在文化影響下，男女性別對健康生活品質的知覺差異或是對健康的注重程度之不同。

除性別之外，年齡通常亦會對SF-36的分數具有影響力。一般而言，身體生理健康領域會隨著年齡的增加而遞減，但心理健康領域則較不受到年齡的影響。在本研究中，研究者以年齡作為效標以檢驗年輕與較老成人之間的差異，分析結果符合原始假設，顯示SF-36台灣版在年齡上的區分效度符合原始心理計量建構的概念。其他國家的常模資料通常亦肯定年齡與身體生理健康領域呈現反比關係，但對於年齡與心理健康之間的關係，

則目前尚無定論。

本研究分析此次國健調查中有關SF-36台灣版的施測資料，並從而建立SF-36台灣版的常模。此常模結果對日後使用SF-36台灣版的研究者相當重要，因其可作為解釋SF-36台灣版施測結果的重要參考依據。常模的建立除有助於SF-36分數的解釋，但更重要的則是未來有關SF-36台灣版的效度研究。由於效度代表心理計量所呈現的分數是否可真正代表其預期欲測量的概念，而本文雖然以年齡作為區分效度的檢驗，但此分析僅能算是初步的結果，日後更有賴台灣的研究者依據不同的研究目的檢測其研究假設，以期增進SF-36台灣版的效度及解釋力。

誌 謝

本研究資料來自民國九十年台灣地區國民健康訪問調查的結果，本研究感謝財團法人國家衛生研究院(計畫編號HP-090-PP-01, HP-090-PP-01, HP-091-PP-01, BS-090-PP-04, BS-091-PP-04)及行政院衛生署國民健康局的經費補助，也感謝所有執行及監督調查的工作人員。

參考文獻

1. Awad AG, Voruganti LN. Intervention research in psychosis: issues related to the assessment of quality of life. *Schizophr Bull* 2000; **26**:557-64.
2. Gandek B, Ware JE, Jr. Methods for validating and norming translations of health status questionnaires: the IQOLA Project approach. *J Clin Epidemiol* 1998; **51**:953-9.
3. Sartorius N. The World Health Organization method for the assessment of health-related quality of life (WHOQOL). In: Walker S, Rosser R, editors. "Quality of Life Assessment: Key Issues in the 1990's." Boston, MA: Kluwer Academic Publishers, 1993;201-7.

4. Ware J, Snow K, Kosinski M, Gandek B. SF-36 Health Survey: Manual & Interpretation Guide. Boston, MA: Nimrod Press; 1993.
5. Gandek B, Ware JE Jr., Aaronson NK, et al. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQOLA Project. *J Clin Epidemiol* 1998;**51**:1149-58.
6. Gandek B, Ware JE, Aaronson NK, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. *J Clin Epidemiol* 1998;**51**:1171-8.
7. Ware JE, Jr., Keller SD, Gandek B, Brazier JE, Sullivan M. Evaluating translations of health status questionnaires. *Int J Technol Assess Health Care* 1995;**11**:525-51.
8. Ren XS, Amick B 3rd, Zhou L, Gandek B. Translation and psychometric evaluation of a Chinese version of the SF-36 Health Survey in the United States. *J Clin Epidemiol* 1998;**51**:1129-38.
9. Lam CL, Gandek B, Ren XS, Chan MS. Tests of scaling assumptions and construct validity of the Chinese (HK) version of the SF-36 Health Survey. *J Clin Epidemiol* 1998;**51**:1139-47.
10. Ju JFR, Tseng HM, Tsai YJ. Assessment of health-related quality of life in Taiwan(I): development and psychometric testing of SF-36 Taiwan version. *Taiwan J Public Health* 2003;**22**:501-11.
11. Fuh JL, Wang SJ, Lu SR, Juang KD, Lee SJ. Psychometric evaluation of a Chinese (Taiwanese) version of the SF-36 health survey amongst middle-aged women from a rural community. *Qual Life Res* 2000;**9**: 675-83.
12. Chie WC, Huang CS, Chen JH, Chang KJ. Measurement of the quality of life during different clinical phases of breast cancer. *J Formos Med Assoc* 1999;**98**:254-60.
13. American Psychological Association. Standards for Educational and Psychological Tests. Washington, DC: American Psychological Association, 1995.
14. Thumboo J, Fong KY, Machin D, et al. A community-based study of scaling assumptions and construct validity of the English (UK) and Chinese (HK) SF-36 in Singapore. *Qual Life Res* 2001;**10**:175-88.
15. Chang D, Chun C, Takeuchi D, Shen H. SF-36 health survey: tests of data quality, scaling assumptions, and reliability in a community sample of Chinese Americans. *Med Care* 2000;**38**:542-8.
16. Scott KM, Tobias MI, Sarfati D, Haslett SJ. SF-36 health survey reliability, validity and norms for New Zealand. *Aust N Z J Public Health* 1999;**23**:401-6.
17. Sullivan M, Karlsson J. The Swedish SF-36 Health Survey III. *J Clin Epidemiol* 1998;**51**:1105-13.
18. Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. *J Clin Epidemiol* 1998;**51**: 1037-44.

