

行政院國家科學委員會專題研究計畫成果報告
簡式 36 健康問卷電腦適性化：與紙筆施測模式之同質性及在臨床醫療照護的實用性之研究 I（第一年）

Examining equivalence and feasibility of computerized adaptive
SF-36 Taiwan version I

計畫編號：NSC 92-2416-H-182-012-

執行期限：92 年 08 月 01 日至 93 年 07 月 31 日

主持人：曾旭民 長庚大學醫務管理系

一、中英摘要

目的：癌症的健康照護中定期收集病患的健康生活品質資料對於照護策略的擬定與品質的提升有相當重要的幫助。目前的 QoL 問卷在臨床實務的使用率並不高，因為其計分與解釋需要耗費相當多的時間，但這些可以運用資訊科技加以解決。本研究主要檢驗電腦化施測的 SF-36 系統在臨床的可用性與其結果與標準紙筆版之間的同質性。

病人與方法：共有 51 位癌症病人同意在看診前後接受電腦與紙筆兩個版本的施測。除了收集 SF-36 的資料外，病人的電腦焦慮程度也用來檢驗是否會影響到電腦施測的結果。

結果：在電腦與紙筆版本所得的 SF-36 的各分量表與及總和分數的有著高度顯著的相關（從 0.87 到 0.98，所有 $p < 0.001$ ）。電腦焦慮的特質與電腦施測的結果沒有顯著的相關。

討論：SF-36 的電腦版在臨床實務的應用顯得相當可行而且病人也大部分接受這種工具。本研究提供了電腦版本與標準紙筆版本同質性，因此對於電腦版本的臨床應用更加可行

關鍵詞：簡式-36, 電腦化測驗, 同質性, 癌症患者生活品質

Abstract

Purpose: Longitudinal health-related quality of life (QoL) data are important for cancer care in routine clinical practice. However, most existing QoL measures were not developed for routine clinical use. Information

technology is making the routine measurement of QoL more feasible. In this study, we evaluate the equivalence and feasibility of a computerized SF-36 in comparison with standard paper mode in a specialist radiotherapy clinics.

Patients and Methods: A total of 51 adult cancer patients were assigned to complete both paper and computer SF-36 versions before and after clinical consultation.

Correlation and paired-t test were applied to examine the equivalence of both administration modes. Personal characteristics such as computer anxiety was also assessed to evaluate its influences upon patients responses under the computerized administration.

Results: There were very significant correlation between the computer and paper version (0.87 to 0.98) for all SF-36 scales and summary scores (all $p < 0.001$). No significant effects were found for the degree of computer anxiety upon responses towards the computerized version.

Discussion: Computerized SF-36 is feasible and acceptable for routine daily clinical use. It provides responses from the computer mode are comparable to those to the paper form, and ensure the use of computerized SF-36 version will not biased by the personal characteristics related to computer anxiety.

Keywords : SF-36, Computerized testing, , Equivalence, Quality of Life of Cancer Patients

二、緣由與目的

在健康權日益受到重視的今日，探討健康相關生活品質（Health-Related Quality of Life, HR-QoL）的研究已吸引更多的研究者致力發展衡量族群健康的方法。一般而言，臨床評估的方法在過去比較侷限於病因及病理現象的改善，健康照護結果（outcome）主要以客觀的生理數據作為指標。近年來則因為強調病人在療程中的主觀感受應為治療結果的重要依據，西方社會日漸重視健康照護結果除了讓病人在生理上舒緩症狀的壓力外，同時也應該能夠帶來病人良好的健康生活品質。較常被使用而且有國際化比較的一般型 HR-QoL 測量工具包括有 SF-36 健康狀態調查表、EuroQoL 歐洲健康生活品質量表、與及 WHOQoL 聯合國健康生活品質量表(Hayes et al., 2000)，這些在台灣都有授權版本而且也有相當的研究證據支持其中文版具有良好的信度與效度。雖然 QoL 可以作為疾病的照護品質的指標，但目前在應用上還是較偏重於新藥檢測或是學術研究上，而非臨床上用來幫助醫生或病人瞭解療程的變化。這些 QoL 測量工具之所以無法在臨床上廣泛被使用的主因在於其計分方式通常相當繁複，無法及時提供臨床醫師相關的資訊以利於疾病的治療。

近年來資訊科技快速的發展使得將 QoL 應用於臨床每日的治療上變得更加可行。心理測驗工具的電腦化不僅可以將施測、計分與及解釋等過程自動化減少臨床人員的負擔同時也可以利用電腦的儲存功能依據個別病人的資料以多媒體（如圖形）的方式加以顯現，幫助臨床專業人員的判讀，而且通常具有高度經濟效益(Lofland, Schaffer, & Goldfarb, 2000)。另外，電腦化心理測驗還可以依據個人在回答問卷的特質進行問卷題本的選擇達到適性化心理測驗(Wainer, 1990)。由於 HR-QoL 測量相當重視個人化與及治療過程的評估，建立

HR-QoL 測量電腦適性化系統使其可以將施測、計分、解釋等過程自動化可以增進在臨床上的高度實用性。在近 10 年來，使用電腦化 HR-QoL 測量的研究與應用日漸增多。這些研究主要集中在長期性疾病病人（如癌症或糖尿病），而其研究目的則包括檢測電腦化 HR-QoL 與傳統紙筆測驗之間的同質性(Fortner, Okon, Schwartzberg, Tauer, & Houts, 2003; Pouwer, Snoek, van der Ploeg, Heine, & Brand, 1998; Velikova et al., 1999)、對治療過程的幫助(Lindblad, Ring, Glimelius, & Hansson, 2002; Lofland et al., 2000; Taenzer et al., 2000; Velikova et al., 2004; Velikova, Brown, Smith, & Selby, 2002)、與及可用性與使用者滿意度(Buxton, White, & Osoba, 1998; Carlson, Specia, Hagen, & Taenzer, 2001; Cull et al., 2001; Taenzer et al., 1997)。

由於電腦化 HR-QoL 能夠給予醫護人員即時回饋與過去資料統整的能力，醫生透過電腦即時計分與自動產生的解釋可以來比較個別病人在不同時間的 HR-QoL 資料，進而能夠更清楚接受化療病人的變化而進行處方的調整。雖然電腦化 QoL 具有相當的優點，但要在臨床上應用必須先確保電腦化版本與原來紙筆版本之間的同質性，如此才能運用過去的紙筆測驗特質於新的電腦版本中，增進該版本解釋的能力。在過去研究中顯示可能會因為受試者與資訊科技有關的特質（如電腦焦慮、電腦經驗、電腦化測驗的接受度等）影響到測驗的結果，這種影響對性格、態度、或症狀知覺等心理測驗特別顯著，因此是否能將紙筆測驗所得的結果直接使用於電腦化測驗的結果目前尚有爭議(George, Lankford, & Wilson, 1992; Lankford, Bell, & Elias, 1994; Tseng, Macleod, & Wright, 1997)。

心理測驗屬性對電腦測驗結果的影響主要是來自於電腦施測具有保密、降低社會預期效應等特性(Wilkerson, Nagao, &

Martin, 2002)。因此若心理測驗屬性為測量認知或速度能力 (cognitive or speed ability) 則較不受到施測媒介的影響(Klinck, 1998; Mead & Drasgow, 1993), 但若為高敏感度的個人資訊或是性格、態度、與及症狀測量等則較容易受到影響(George et al., 1992; Honaker, 1988; Tseng, Tiplady, Macleod, & Wright, 1998)。除了測驗屬性以外, 文獻研究中也嘗試瞭解受試者個人與資訊科技有關的特質是否造成電腦測驗結果的差異。其中較常被提及的為電腦焦慮 (computer anxiety) 特質, 通常具有電腦焦慮特質的受試者在電腦化心理測驗中的表現比較容易受到焦慮情緒的影響, 而導致測驗表現上的差異(Gaudron, 1998; Lankford et al., 1994; Tseng et al., 1998)。由於 QoL 測量主要和個人的臨床症狀有關, 在內容上有些東西也牽涉到較為隱私的部分 (例如性行為的感受等), 因此電腦化測驗版本與紙筆版本的同質性研究更顯的重要。文獻中關於 QoL 的同質性研究尚未取得共識的結論, Velikova et al. (1999)研究顯示在歐洲版癌症治療生活品質問卷 (EORTC QLQ-C30) 電腦化版本上有關負向情緒 (如疲倦) 的反應分數較高, 而其它題本則沒有顯著差別。不過在 Pouwer (1998) 使用 Well-being Questionnaire (WBQ) 在糖尿病人與及 Wilson et al. (2002) 使用 SF-36 於癌症病人上, 發現兩種施測模式所得結果呈現高度的相關。

本研究主要的目的在於建立一般型 HR-QoL 中常被使用的 SF-36 中文版的電腦適性化測量資訊系統, 以利於臨床與研究上的使用。在第一年的研究主要瞭解 SF-36 電腦化測量系統與原來紙筆測驗結果的同質性 (equivalence)。藉由此研究不僅可以讓國內研究者與醫療臨床實務者有一套自動化的 SF-36 施測與解釋系統可以簡化其施測與結果分析的負擔, 同時也可以發展出本土化 SF-36 的總合分數計算公式, 並與國際間進行比較。

三、結果與討論

本年度的研究目的主要在於瞭解 SF-36 電腦化與紙筆版本之間的同質性, 在研究設計上採用重複測量的方法, 受試者在醫院看診前後同時接受兩個版本的測量。共有 59 名頭頸部癌症病人自願接受測試, 但由於看診與時間的因素, 最後有 51 位病人完整完成兩個版本的 SF-36 測試。這些病人超過 9 成為男性 (N=46), 年齡分佈的平均值為 50.7 (SD=12.3) 歲。

研究所使用的工具包括有台灣 SF-36 第一版與及電腦焦慮問卷。SF-36 包含 36 項問題組成 8 個量表, 主要為擷取自 Medical Outcomes Study(MOS)的健康生活品質問卷 (Steward & Ware, 1992)。身心健康狀態的 8 個概念分別為身體生理功能 (physical functioning, PF)、因身體生理問題角色受限 (role limitation due to physical problems, RP)、身體疼痛 (bodily pain, BP)、一般健康狀況 (general health, GH)、活力 (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)、及心理健康 (mental health, MH), 此外尚含一項自評健康變化 (reported health transition)。目前台灣版已經建有常模資料而且其題本信度、常模與及初步效度檢驗顯示 SF-36 台灣版與原版本之間具有良好並可與國際相互比較的心理計量特質 (Lu, Tseng, & Tsai, 2003; Tseng, Lu, & Tsai, 2003)。電腦焦慮問卷為 13 題之問卷, 可分為兩項重要因素: 情緒反應和認知反應。整體問卷的內部一致性 (Cronbach α) 為 0.76, 而分測驗中情緒反應與認知反應兩個分測驗的 Cronbach α 都為 0.89。

施測的程序為隨機分配受試者於電腦-紙筆版本 (N=24) 或紙筆-電腦版本 (N=27) 的兩個組別中, 以避免學習效應。在電腦-紙筆版本組別中, 受測者先接受電腦化 SF-36 版本, 在經過看診之後, 再接受紙筆版本的測驗, 而紙筆-電腦版本組別的施測

版本順序則相反。

檢測電腦與紙筆版的差異主要從兩種施測版本所得的 SF-36 分數的相關與及 paired-t 檢定來加以檢測。如 Table 1 所示，在兩種版本的 SF-36 的八個量表與身體總和及心理總和分數的相關係數，顯示兩種施測方法所得的結果具有極高的相關 (Pearson's r 介於 0.87 與 0.98 之間，所有的相關係數 $p < .001$)。另外利用重複施測的 t 檢定，也顯示兩者之間所得的平均值相當接近，並沒有統計上顯著的差異 (如 Table 2)。另外病人個人電腦焦慮特質是否會影響其在填答電腦施測上的分數，主要以焦慮特質分別與電腦施測與紙筆施測的 SF-36 分數比較其相關係數，結果顯示施測模式與電腦特質並沒有顯著的相關。這些結果顯示利用電腦來施測可以得到與原來問卷施測上非常類似的結果，顯示兩者可以作為平行的施測模式。本研究的結果顯示電腦版與紙筆版的 SF-36 具有良好的平行性，此結果與多數的電腦化 QoL 的研究結果相似 (Buxton et al., 1998; Fortner et al., 2003; Lofland et al., 2000; Velikova et al., 1999; Wilson et al., 2002)。而且因為使用筆觸式的平板電腦，使得電腦版的 SF-36 更接近原使紙筆版本的施測模式，所以過去文獻中常見的電腦焦慮特質對電腦化施測並沒有顯著的影響。

四、計畫成果自評

本年度的研究主要著重於 SF-36 電腦化施測系統的發展與及檢驗其與原紙筆版本之間的同質性。本年度的研究結果基本上符合原本實驗設計的預期，顯示電腦版本與紙筆版本之間基本上是相容並行的，而且在電腦施測方面，由於運用筆觸式的 tabloid 電腦，由於其介面與原紙筆版本相當接近，因此病人電腦焦慮的特質的影響並不顯著。今年度的研究結果可以順利的加以運用進行下年度電腦適性化的測驗系統的發展，同時也確立了原有的紙筆 SF-36 所建構的資料與

測驗的相關計量指標可以運用於 SF-36 的電腦版本上。

在研究過程中，較為棘手的部分在於病人的配合度上。由於採用的是重複測量的研究方法，病人需要在看診之後仍然回到施測室接受另一版本的測試，還好兩個版本所花的時間基本上不多，但沒有長庚醫院林口放射科醫師們的協助與及病人的容忍與支持，本研究根本不可能完成，在此特別感謝。

由於醫學的進步許多的疾病已經不再是致死性的，而反而常變成為需要長期性的照護。在醫療照護結果不僅需要考量症狀的抒解同時也應思考病人在治療過程中的生活品質之增進。本研究結果利用電腦化 SF-36 台灣版以增進其在臨床實用的價值，可以協助醫護人員在照護病人時能夠有病人的健康生活品質資料作為治療過程決策的參考，同時分析常模資料來得到 SF-36 計算的公式與及個別化施測的方法，也有助於台灣與國際間使用 SF-36 的比較。

五、參考文獻

- Buxton, J., White, M., & Osoba, D. (1998). Patients' experiences using a computerized program with a touch-sensitive video monitor for the assessment of health-related quality of life. *Quality of Life Research*, 7(6), 513-519.
- Carlson, L. E., Specia, M., Hagen, N., & Taenzer, P. (2001). Computerized quality-of-life screening in a cancer pain clinic. *Journal of Palliative Care*, 17(1), 46-52.
- Cull, A., Gould, A., House, A., Smith, A., Strong, V., Velikova, G., et al. (2001). Validating automated screening for psychological distress by means of computer touchscreens for use in routine oncology practice. *British Journal of Cancer*, 85(12),

1842-1849.

- Fortner, B., Okon, T., Schwartzberg, L., Tauer, K., & Houts, A. C. (2003). The cancer care monitor: Psychometric content evaluation and pilot testing of a computer administered system for symptom screening and quality of life in adult cancer patients. *J Pain Symptom Manage*, 26(6), 1077-1092.
- Gaudron, J. P. (1998). The effects of computer anxiety and past computer experience on computerized test performance. *Travail Humain*, 61(3), 263-280.
- George, C. E., Lankford, J. S., & Wilson, S. E. (1992). The effects of computerized versus paper-and-pencil administration on measures of negative affect. *Computers in Human Behavior*, 8(2-3), 203-209.
- Hayes, J. A., Black, N. A., Jenkinson, C., Young, J. D., Rowan, K. M., Daly, K., et al. (2000). Outcome measures for adult critical care: A systematic review. *Health Technology Assessment*, 4(24), 1-111.
- Honaker, L. M. (1988). The equivalency of computerized and conventional mmpi administration: A critical review. *Clinical Psychology Review*, 8(6), 561-577.
- Klinck, D. (1998). Paper-and-pencil versus computerized administration of cognitive ability tests: A study addressing the question of equivalence. *Diagnostica*, 44(2), 61-70.
- Lankford, J. S., Bell, R. W., & Elias, J. W. (1994). Computerized versus standard personality measures: Equivalency, computer anxiety, and gender differences. *Computers in Human Behavior*, 10(4), 497-510.
- Lindblad, A. K., Ring, L., Glimelius, B., & Hansson, M. G. (2002). Focus on the individual--quality of life assessments in oncology. *Acta Oncologica*, 41(6), 507-516.
- Lofland, J. H., Schaffer, M., & Goldfarb, N. (2000). Evaluating health-related quality of life: Cost comparison of computerized touch-screen technology and traditional paper systems. *Pharmacotherapy*, 20(11), 1390-1395.
- Lu, J.-F. R., Tseng, H.-M., & Tsai, Y.-J. (2003). Assessment of health-related quality of life in taiwan (i):Development and psychometric testing of sf-36 taiwan version. *Taiwan J of Public Health*, 22, 501-511.
- Mead, A. D., & Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive ability tests: A meta-analysis. *Psychological Bulletin*, 114(3), 449-458.
- Pouwer, F., Snoek, F. J., van der Ploeg, H. M., Heine, R. J., & Brand, A. N. (1998). A comparison of the standard and the computerized versions of the well-being questionnaire (wbq) and the diabetes treatment satisfaction questionnaire (dtsq). *Quality of Life Research*, 7(1), 33-38.
- Steward, A., & Ware, J. (1992). *Measuring functioning and well-being: The medical outcomes study approach*. Durham, NC: Duke University Press.
- Taenzer, P., Bultz, B. D., Carlson, L. E., Specia, M., DeGagne, T., Olson, K., et al. (2000). Impact of computerized quality of life screening on physician behaviour and patient satisfaction in lung cancer outpatients. *Psycho Oncology*, 9(3), 203-213.
- Taenzer, P. A., Specia, M., Atkinson, M. J., Bultz, B. D., Page, S., Harasym, P., et al. (1997). Computerized quality-of-life screening in an oncology clinic. *Cancer Practice*, 5(3), 168-175.
- Tseng, H.-M., Lu, J.-F. R., & Tsai, Y.-J.

- (2003). Assessment of health-related quality of life (ii): Norming and validation of sf-36 taiwan version. *Taiwan J of Public Health*, 22, 512-518.
- Tseng, H. M., Macleod, H. A., & Wright, P. (1997). Computer anxiety and measurement of mood change. *Computers in Human Behavior*, 13(3), 305-316 (SSCI).
- Tseng, H. M., Tiplady, B., Macleod, H. A., & Wright, P. (1998). Computer anxiety: A comparison of pen-based personal digital assistants, conventional computer and paper assessment of mood and performance. *British Journal of Psychology*, 89(Part 4), 599-610.
- Velikova, G., Booth, L., Smith, A. B., Brown, P. M., Lynch, P., Brown, J. M., et al. (2004). Measuring quality of life in routine oncology practice improves communication and patient well-being: A randomized controlled trial. *J Clin Oncol*, 22(4), 714-724.
- Velikova, G., Brown, J. M., Smith, A. B., & Selby, P. J. (2002). Computer-based quality of life questionnaires may contribute to doctor-patient interactions in oncology. *British Journal of Cancer*, 86(1), 51-59.
- Velikova, G., Wright, E. P., Smith, A. B., Cull, A., Gould, A., Forman, D., et al. (1999). Automated collection of quality-of-life data: A comparison of paper and computer touch-screen questionnaires. *Journal of Clinical Oncology*, 17(3), 998-1007.
- Wainer, H. (1990). *Computerized adaptive testing: A primer*. Hillsdale, NJ: Lawrence Erlbaum Associations.
- Wilkerson, J. M., Nagao, D. H., & Martin, C. L. (2002). Socially desirable responding in computerized questionnaires: When questionnaire purpose matters more than the mode. *Journal of Applied Social Psychology*, 32(3), 544-559.
- Wilson, A. S., Kitas, G. D., Carruthers, D. M., Reay, C., Skan, J., Harris, S., et al. (2002). Computerized information-gathering in specialist rheumatology clinics: An initial evaluation of an electronic version of the short form 36. *Rheumatology (Oxford)*, 41(3), 268-273.

六、圖表

Table 1. Pearson's correlation coefficients between paper and computer version of SF-36 scores

	SF-36 Scales								Summary	
	Physical	Role	Bodily	General	Vitality	Social	Role	Mental	PCS	MCS
	Function	Physical	Pain	Health		Function	Emotional	Health		
Study Group	.98	.97	.94	.91	.87	.92	.90	.91	.96	.94

Data are Pearson's correlation coefficients (n=51); $p < 0.001$ for all coefficients.

PCS=Physical Component Summary; MCS=Mental Component Summary

Table 2. Comparison of score distribution between computer and paper versions of the SF-36

	Administration Mode	
	Computer	Paper
SF-36 Scale	Mean (S.D.)	Mean (S.D.)
Physical Function	82.6 (13.4)	82.4 (13.7)
Role Physical	50.0 (41.1)	47.6 (40.2)
Bodily Pain	73.4 (16.9)	72.6 (15.8)
General Health	51.6 (8.4)	51.5 (8.5)
Vitality	58.3 (16.3)	60.5 (20.0)
Social Function	67.3 (28.9)	69.6 (28.9)
Role Emotional	38.1 (42.5)	42.1 (39.3)
Mental Health	62.1 (21.7)	62.7 (22.8)
Summary Scales		
Physical Component Summary	43.6 (6.1)	43.0 (7.2)
Mental Component Summary	42.0 (12.6)	43.2 (14.9)

All computer versus paper modes comparisons are *not* significant as using paired-t test