

國立臺灣大學公共衛生學院流行病學與預防醫學研究所

碩士論文

Graduate Institute of Epidemiology and Preventive Medicine


College of Public Health

National Taiwan University

Master Thesis

中文化兒童排尿障礙問卷量表之信度及效度統計分析

Statistical Analysis of Reliability and Validity of a Chinese Version  
of Dysfunctional Voiding Symptom Score (DVSS) Questionnaire



張尚仁

**Shang-Jen Chang**

指導教授：戴政 教授  
陳秀熙 教授

Supervisors: Professor John Jen Tai, Ph.D

Professor Tony Hsiu-Hsi Chen, Ph.D.

中華民國一〇〇年八月

August, 2011

## 中文摘要

兒童排尿障礙為一個多重病因的疾病，其臨床的症狀表現：包含了日間尿失禁，急尿，急尿屈膝，排尿困難及便秘等症狀。從西元 2000 年起，至少有四種兒童排尿障礙問卷量表被提出來測量兒童排尿障礙。但至今仍沒有中文化的量表被提出，因此我們便著手翻譯中文問卷量表，並以適當統計方法分析信度及效度。本研究設計為年齡及性別配對之個案對照研究。個案組為從西元 2008 九月開始，在門診收集 60 位（38 位女生，22 位男生）經由臨床評估及尿路動力學診斷為排尿障礙的兒童，予其父母或主要照顧者填寫排尿障礙問卷，並於一週內填寫第二次量表。我們收集了基本資料，包含了身高，體重，尿液分析，尿路動力學檢查結果。另外於幼稚園及小學收集年齡及性別相同的健康兒童作為對照組，同樣予其父母或主要照顧者填寫排尿障礙問卷，共 235 位（90 位男生及 145 位女生）作為對照組。

原版兒童排尿障礙問卷量表(DVSS)為英文，由 Farhat 等人提出，包含了十個問題，每題權重相同為 0 到 3 分。其中七個問題與排尿症狀相關，2 個與便秘相關，一個與壓力事件相關。中文化問卷由楊緒棟翻譯，並再翻譯成英文。中文版問卷並由台灣泌尿科醫學會五位小兒泌尿委員會委員審視內容效度。

信度分析方面共有 2 項：1. 內部一致性以 Cronbach's alpha 檢驗，2. 重測信度以組內相關係數 (intraclass correlation) 檢驗。我們採用羅吉斯迴歸及 ROC 曲線分析來作預測效度分析。此外，並以因素分析，將十個問題分類，得到不同的建構面向。我們並採用貝氏定理來分析在不同盛行率地區作篩檢時，同樣量表的分數可能患有排尿障礙的機率及風險比。

本研究個案平均總分為  $9.65 \pm 3.87$  分，而對照組為  $4.13 \pm 2.60$  分。個案量表的量表內部項目相關係數 (Cronbach's alpha) 為 0.448，顯示此量表所測量的疾病症狀為異質化的組成。以組內相關係數評估重測信度為 0.89。經由 ROC 曲線分析決定量表總分切點為 6.66，其 AUC

為 0.89，敏感性為 81.67%，而特異性為 82.63%。因數分析發現所有的症狀可以分類為三種建構：1. 尿失禁及急尿症狀，2. 排泄困難症狀，3. 疼痛及壓力症狀。依貝氏方法調整以 DVSS 於不同盛行率地區，或於小兒泌尿門診篩檢時的預測機率，同樣的 DVSS 總分，在不同地區為個案的機率不同，比如同樣的 DVSS 總分 8 分，在社區（盛行率 2%）可能為個案的機率只有 8%，風險比為 0.09，而在門診（盛行率 30%）則機率增加至 63%，風險比為 1.74。

根據本研究結果，中文化兒童排尿障礙問卷量表具有良好的重測信度，並且可以用來預測個案的效度，由貝氏方法發現中文化兒童排尿障礙問卷應用於社區及門診篩檢時，同樣的分數會有不同的風險比。



關鍵字：兒童排尿障礙、問卷量表、信度、效度、因素分析、貝氏定理

## ABSTRACT

Dysfunctional voiding in children is a multifactorial disease presenting with a variety of symptoms including daytime incontinence, urgency, curtsyng, difficult voiding and constipation. In spite of four different scales have been validated to measure pediatric dysfunctional voiding since 2000, there is still lacking of psychometrically-sound scales in Chinese to measure pediatric dysfunctional voiding, we performed an age- gender-matched case-control study to evaluate the reliability and validity of Dysfunctional Voiding Symptom Score (DVSS) in Chinese version. From Sep 2008 to Feb 2011, we enrolled 60 children (38 girls and 22 boys, case) with diagnosis of dysfunctional voiding by clinical evaluation and urodynamic studies in clinics of pediatric urology. Baseline characteristics were recorded including age, gender, body height, weight, urine analysis and parameters associated with urodynamic studies. The DVSS was completed at clinics by the parents who mainly take care of the child in the presence of the child. After one week, the DVSS was administered at home and returned to clinics for evaluation of test-retest reliability. For the control group, we enrolled 235 age- and gender- matched healthy children from kindergartens and elementary schools.

The DVSS consisted of 10 items with the score of each item scores ranging from 0 to 3. Among these items, seven were related to voiding, two were related to defecation and one was related to stressful events. The DVSS was translated into Chinese and did back translation into English. The DVSS in Chinese version was reviewed by five members in the committee of pediatric urology of Taiwan Urological Association with respect

to content validity.

The reliability of DVSS questionnaires completed by cases was assessed with Cronbach's alpha test (internal consistency) and intraclass correlation (ICC, test-retest reliability). The predictive validity was first analyzed with logistic regression and was further assessed by Receiver Operating Characteristic curve analysis. Factor analysis was used to classify the ten symptoms into latent factors and estimate factor score of individual latent variables. Finally, we applied Bayesian method to evaluating the posterior probability and odds of dysfunctional voiding in children screened with DVSS from areas of different prevalence rates.

Mean total scores of DVSS in case and control group were  $9.65(\pm 3.87)$  and  $4.13(\pm 2.60)$ , respectively. The Cronbach's alpha coefficient was 0.448 for the ten items among DVSS indicating a heterogeneous composition of symptoms in dysfunctional voiding. Intraclass correlation coefficient for test-retest reliability was 0.89. The cut-off for total DVSS score was 6.66, yielding 81.67% of sensitivity and 82.63% of specificity. Factor analysis revealed three latent variables (overactive bladder symptoms, difficult elimination symptoms, stress and dysuria symptoms). Using Bayesian method, the applications of DVSS in areas of different prevalence figures give an indication of significantly different probabilities of dysfunctional voiding, 8% and 63% in a community with 2 % prevalence and pediatric urology clinics with 30 % prevalence, respectively give a total score equal to 8.

Conclusions: DVSS in Chinese version is reliable and valid to predict children with dysfunctional voiding. Through Bayesian method, we found significantly different

probabilities and odds of having dysfunction voiding in children with the same total DVSS scores from areas of different prevalence figures.

Keywords: pediatric dysfunctional voiding, questionnaire scale, reliability, validity, factor analysis, Bayesian theorem



# CONTENTS

中文摘要.....	i
ABSTRACT.....	iii
CONTENTS.....	vi
List of figures.....	ix
List of tables.....	x
<b>Chapter 1 前言.....</b>	<b>1</b>
<b>Chapter 2 文獻回顧.....</b>	<b>3</b>
2.1 兒童排尿障礙及四種相關問卷量表.....	3
2.1.1 Dysfunctional Voiding Symptom Score .....	4
2.1.2 Dysfunctional voiding and incontinence scoring system .....	4
2.1.3 Assessment of urinary symptoms in children with dysfunctional elimination syndrome .....	5
2.1.4 Symptom Score for Dysfunctional Elimination Syndrome.....	5
2.2 問卷翻譯.....	9
2.3 問卷題目分數的比重及加權.....	10
2.4 問卷信度的評估.....	11
2.4.1 量表之內項目及問題的一致性 .....	12
2.4.2 組內相關係數 (intraclass correlation, ICC) .....	13
2.5 量表的效度評估.....	17
2.6 探索性因素分析 (exploratory factor analysis) .....	20
<b>Chapter 3 研究方法.....</b>	<b>26</b>
3.1 研究設計.....	26

3.2	測量問卷翻譯及內容效度.....	26
3.3	問卷信度分析.....	28
3.4	統計模型建立.....	28
3.4.1	信度測試.....	28
3.4.2	效度建立：.....	29
3.4.3	因素分析（factor analysis）.....	31
3.4.4	運用貝氏方法來調整DVSS的預測機率。.....	33
<b>Chapter 4</b>	<b>結果.....</b>	<b>36</b>
4.1	基本資料描述性統計.....	36
4.2	遺漏值處理.....	40
4.3	信度分析.....	41
4.3.1	重測信度（Test-retest reliability）.....	41
4.3.2	內部一致性（internal consistency）.....	42
4.4	預測效度分析（predictive validity）.....	43
4.4.1	切點選擇.....	43
4.5	主因素分析(principal factor analysis）.....	47
4.5.1	依個案症狀作主因素分析（Principal factor analysis）....	47
4.5.2	以對照組兒童的症狀作主因素分析（Principal factor analysis）	50
4.5.3	以所有個案及對照組症狀作主因素分析（Principal factor analysis）	53
4.6	羅吉斯迴歸模型之AUC值及模型適切度.....	56
4.7	貝氏定理應用於以DVSS篩檢社區及門診兒童.....	60
<b>Chapter 5</b>	<b>討論.....</b>	<b>61</b>





## List of figures

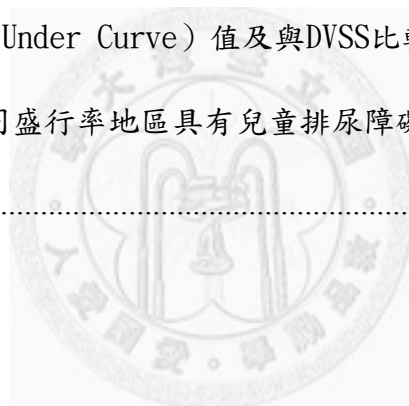
Figure 2.5-1 Receiver Operating Characteristic curve (ROC curve) .....	20
Figure 2.6-1 scree diagram .....	24
Figure 3.4-1 選擇不同的切點 (cut-off point) 來分辨個案及非個案.....	29
Figure 4.4-1 建立預測效度的ROC曲線圖形 (所有個案) .....	44
Figure 4.4-2 以羅吉斯迴歸建立預測效度的ROC曲線圖形 (女童) .....	45
Figure 4.4-3 以羅吉斯迴歸建立預測效度的ROC曲線圖形 (男童) .....	46



## List of tables

Table 2.1-1 兒童排尿障礙問卷量表的相關文獻回顧及結果整理 .....	7
Table 4.1-1 個案及對照組的基本資料分佈與比較 .....	37
Table 4.1-2 女童個案及女童對照組的基本資料分佈與比較 .....	38
Table 4.1-3 男童個案及男童對照組的基本資料分佈與比較 .....	39
Table 4.2-1 問題一至問題十的遺漏值數量 .....	40
Table 4.3-1 個別問卷項目及總分的重測信度（組內相關係數） .....	41
Table 4.3-2 個案及對照組項目的內部一致性（internal consistency）之Cronbach's alpha值 .....	42
Table 4.4-1 選擇不同總分切點的敏感性及特異性（所有個案） .....	44
Table 4.4-2 選擇不同總分切點的敏感性及特異性（女童） .....	45
Table 4.4-3 選擇不同總分切點的敏感性及特異性（男童） .....	46
Table 4.5-1 依個案症狀作因素分析之因素特徵值（個案組）及其解釋共同性比例	47
Table 4.5-2 個案症狀因素分析的原始因素負荷，共同性及特異變異（因素負荷>0.3, *）	48
Table 4.5-3 個案症狀因素分析，經過Varimax軸線旋轉後的因素負荷，共同性及特異變異（因素 負荷>0.3, *）.....	49
Table 4.5-4 以對照組兒童的症狀因素分析之因素特徵值（對照組）及其解釋共同性比例	50
Table 4.5-5 對照組症狀因素分析的原始因素負荷，共同性及特異變異（因素負荷>0.3, *） .....	51
Table 4.5-6 對照組症狀因素分析，經過Varimax軸線旋轉後的因素負荷，共同性及特異變異（因 素負荷>0.3, *）.....	52

Table 4.5-7 所有個案及對照組症狀因素分析之因素特徵值（對照組）及其解釋共同性比例 .....	53
Table 4.5-8 所有個案及對照組症狀作因素分析的原始因素負荷，共同性及特異變異(因素負荷 >0.3, *).....	54
Table 4.5-9 所有個案及對照組症狀作因素分析，經過Varimax軸線旋轉後的因素負荷，共同性 及特異變異（因素負荷>0.3, *）.....	55
Table 4.6-1 各個模型Receiver Operating Characteristic（ROC）curve 的AUC（Area Under Curve）值，模型適切度(Akaike' s information criterion)，變數迴歸係數及變 數風險比之比較。.....	58
Table 4.6-2 各個模型AUC（Area Under Curve）值及與DVSS比較 .....	59
Table 4.7-1 各個DVSS總分在不同盛行率地區具有兒童排尿障礙的機率（probability）及風險 比（odds ratio）.....	60



## Chapter 1 前言

兒童排尿障礙(dysfunctional voiding, 簡稱DV)的診斷為當患者之神經學檢查正常, 卻有膀胱逼尿肌過動, 排尿時骨盆底肌肉及括約肌無法完全放鬆或活性過高, 造成排尿時膀胱出口阻塞, 導致有尿失禁, 急尿, 小便斷斷續續, 排尿困難, 便秘、尿路感染, 和膀胱輸尿管尿液逆流等症狀及問題[1]。國際小兒尿失禁醫學會對於兒童排尿障礙的診斷共識認為診斷排尿障礙, 除了醫師臨床症狀評估及身體檢查之外, 尚需要非侵襲性的尿流速檢查, 肌電圖檢查及膀胱殘尿量檢查來確立診斷, 並非全部的小孩都要施行侵襲性高的全套尿路動力學檢查[1]。因此, 診斷的準則除了客觀的尿路動力學檢查結果之外, 臨床醫師也需要一個量表來系統性的評估病童所有可能的相關症狀及表現, 並量化各項症狀的嚴重程度。就臨床治療觀點而言, 此量表除了可以運用於門診評估就診兒童是否有患有排尿障礙的疾病之外, 更可以用來追蹤測量患有排尿障礙病童接受治療之後的症狀及病情變化, 並測量比較不同治療方法對於治療病人的效果。就預防觀點而言, 也可以利用此問卷量表於社區兒童, 預測受檢兒童是否具有排尿障礙的疾病, 以利早期治療。

自西元 2000 年開始, 其他國家的小兒泌尿科醫師已經發展四種具有信度及效度的英文及土耳其文問卷量表[2-5], 但台灣並沒有具有信度及效度的中文化兒童排尿障礙問卷量表可供臨床使用。因此, 本研究目的為中文化兒童排尿障礙問卷(Chinese version of Dysfunctional Voiding Symptom Score, DVSS, Farhat et al 2000[2]) 信度(reliability)與效度(validity)之分析和探討。此外, 由於兒童排尿障礙患者合併有尿失禁, 頻尿, 急尿, 便秘等相關症狀, 其所定義之症狀變項間可能會有高度相關, 因此本研究另一目的為如何利用相關之統計方法來評估建構兒童排尿障礙症狀的潛在因子, 並將所有的症狀變項依這些潛在因子進行分類。最後, 運用中文化兒童排尿障礙問卷在不同盛行率地區作篩檢時, 比如在小兒泌尿科門診或學校, 同樣中文化兒童排尿障

礙問卷總分的兒童，其患有排尿障礙的風險可能並不相同，如何運用相關統計方法來處理及調整中文化兒童排尿障礙問卷在不同盛行率地區進行篩檢時的預測機率，亦為本研究的目的之一。



## Chapter 2 文獻回顧

### 2.1 兒童排尿障礙及四種相關問卷量表

兒童排尿障礙為一種排尿時膀胱及括約肌失調造成的一種疾病,合併有尿失禁,頻尿,急尿,便秘等症狀[6, 7]。一般認為排尿障礙的病態生理學為多重因素,有可能是因為排尿學習的障礙,神經成熟遲緩或者基因遺傳之故。患有此疾病的兒童會有膀胱逼尿肌過動,使得兒童為了避免尿失禁,會刻意或不自主的收縮尿道括約肌及骨盆底肌肉,造成斷斷續續的尿流,並因為膀胱壓力上升及出口阻塞,造成尿液逆流,尿液殘留和尿路感染。這些兒童除了尿路的症狀,常常伴隨有便秘的症狀,而便秘的解除,往往也可以減少排尿相關的症狀,其原因可能為膀胱和直腸有相同的胚胎來源,神經支配及相近的解剖結構[8]。

由於各個研究的診斷標準不一,因此並不知道確切的盛行率,只能靠一些小型的研究認定一般社區兒童的排尿障礙盛行率約為 2%至 15%[1, 7, 9],至於小兒泌尿門診來求診的兒童,具有排尿障礙的盛行率約為 30%[10]。在臨床面對有排尿障礙的兒童時,小兒泌尿科醫師除了病史及症狀詢問,身體評估之外,需要以非侵襲性的尿路動力學檢查來診斷兒童是否有排尿障礙。全套尿路動力學檢查因為具有侵入性,會有疼痛及尿路感染的副作用,因此國際小兒尿失禁醫學會的診斷共識已經不建議常規使用全套的尿路動力學檢查來診斷這類的病人,主要以臨床症狀及非侵襲性的尿流速,肌電圖及殘尿超音波為診斷標準[1]。因此,小兒泌尿科醫師便需要設計一個非侵入性評估問卷量表,系統性評估兒童的症狀,並預估門診來診的兒童是否患有排尿障礙,並以此作為是否需要作進一步追蹤小孩子症狀的指標,此外也可以用於社區作為篩檢兒童排尿障礙之用。

本論文經文獻回顧後找到的四種大同小異的排尿障礙問卷摘要於 Table 2.1-1。

### 2.1.1 Dysfunctional Voiding Symptom Score

西元 2000 年, Farhat等人[2]提出第一個兒童排尿障礙問卷 (Dysfunctional Voiding Symptom Score, DVSS), 其研究設計為個案對照研究法, 個案組為經臨床診斷患有排尿障礙的兒童, 對照組則為從社區及學校收集相同年紀的兒童, 兩組兒童都填寫DVSS問卷。其問卷項目設計為修改自成人的排尿症狀問卷, 共有十個問題, 其中七個與排尿相關, 兩個與大便相關, 一個與壓力事件相關。每個問題權重相同, 為 0-3 分, 若以Receiver Operating Characteristics curve分析來選擇切點, 發現男生在 9 分以上, 而女生在六分以上為有排尿障礙的切點。其缺點為在設計問卷項目時, 沒有專家會議及病童訪談 (focus group interview), 並詢問病童哪些症狀與其疾病較為相關, 或者有不易回答之處。此外亦沒有分析問卷內部一致性及重測信度。其他學者亦批評其問卷項目缺少了頻尿及大便失禁這兩個項目。而且原版的問卷第五題語意不清, 原題目為詢問小孩是否一天只有去廁所一至兩次, 應該是詢問小孩子一天是否只有尿尿一至兩次。

### 2.1.2 Dysfunctional voiding and incontinence scoring system

西元 2005 年, Akbal等人[3]發展了一套土耳其語的兒童排尿障礙問卷, 研究設計一樣為個案對照研究法, 個案組為有在門診診斷患有排尿障礙的兒童, 另外一組為從社區及學校收集相同年紀的健康兒童。其問卷項目設計包含 14 個問題, 12 個與排尿相關, 1 個與大便相關, 一個與生活品質相關。每個問題權重不同, 有兩個選項(是或否)也有三個選項(依嚴重程度分級), 由羅吉斯迴歸的迴歸係數加權決定分數, 總分 0 到 35 分, 以Receiver Operating Characteristics curve分析來選擇切點, 以 8.5 分作切點可有 90%的sensitivity和 90%的specificity, 男女生之間



並沒有差別。其缺點一樣為缺乏專家會議及病童訪談（focus group interview），重測信度和內部一致性。此外，以不同迴歸係數加權決定每個題目分數，容易造成計算困難。最後，有些題目回答為是或否，而有些項目有三個回答選項容易造成答題者混淆。

### 2.1.3 Assessment of urinary symptoms in children with dysfunctional elimination syndrome

西元 2007 年，另外一組學者發表了另一套土耳其語的兒童排尿障礙問卷量表[4]，一樣採取個案對照研究，他們收集了 81 個有排尿障礙的兒童，及 102 位健康對照組。原始問卷一開始先設計了 35 個問題，經由Receiver Operating Characteristics curve分析，剔除Area Under Curve小於 0.6 的問題，最後剩下 15 問題，12 個與排尿相關，2 個與大便相關，一個與壓力事件相關。結果總分為 0-35 分，以ROC分析來選擇切點，發現男生在 4.5 分以上女生 7.5 分以上為有排尿障礙的切點。其缺點為缺乏重測信度及內部一致性分析。此外，將Area Under Curve小於 0.6 的題目剔除，可能會減損了問卷內容效度及建構效度。

### 2.1.4 Symptom Score for Dysfunctional Elimination Syndrome

在 2009 年，Afshar等人[5]發表了另一套問卷，採個案對照研究法，收集了 62 位有排尿障礙的兒童，於門診填寫問卷之後，一週內再填寫第二次問卷，以檢測重測信度，另外在耳鼻喉科門診收集 50 個沒有泌尿道症狀的對照組兒童，以此對照組兒童的排尿障礙分數來測試此問卷量表的效度。其問卷設計為根據之前的文獻探討及專家意見，經過目標族群（focus group）的訪談，將原有的十五個問題剔除了兩個問題（因為問題模稜兩可，或者難以理解），提出了一個十三個問題的量表，以 5-point Likert scale計分，每個問題分數權重一致（weighted equally），0 到 4 分。結果以Receiver Operating Characteristics curve分析發現以 11 分作為切點來分辨個

案，可以有 80%的敏感性及 91%特異性。以Pearson correlation作分析，發現重測信度之相關係數為 0.845。量表內部信度(Cronbach's alpha)為 0.445。因為內部信度較差，所以他們採取因素分析方法分析此量表，發現這十三個問題可以分為四個面向（incontinence, urgency, obstructive symptoms, and fecal complaints），不過他們並沒有將這些潛在因素進一步分析，並提出哪些症狀歸類至哪一個潛在面向，也沒有就男女之間是否有差異作分析。



Table 2.1-1 兒童排尿障礙問卷量表的相關文獻回顧及結果整理

Author	Study design	Questionnaire	Statistical method	Results	Comments
The dysfunctional voiding scoring system: quantitative standardization of <b>dysfunctional voiding symptom score (DVSS)</b> Farhat W. (2000)[2]	1. Age-matched case-controlled study 2. Case: children with dysfunctional voiding at clinics Control: children without dysfunctional voiding from schools were enrolled for comparison	1. 10-item questionnaire (7 related to voidings, 2 related to bowel habits, 1 related to stressful events, each score ranging from 0 to 3, with total score of 30)	1. ROC curves analysis	1. Boys with DVSS of >9 point (sensitivity 80.95%, specificity 91.30%).and girls with >6 points selected as cut-off point for dysfunctional voiding (sensitivity 92.77%, specificity 87.09%) 2. girls tended to have dysfunctional voiding with odds of 2.9	1. Original form of the DVSS in our study 2. Lacking of expert committee, focus group interview 3. Lacking of reliability and internal consistency analysis 4. Content validity : lacking of stool incontinence, and urinary frequency
<b>Dysfunctional voiding and incontinence scoring system (DVISS)</b> :quantitative evaluation of incontinence symptoms in pediatric population Akbal 2005[3]	1. Age-matched case-controlled study 2. Children with dysfunctional voiding at clinics and children without dysfunctional voiding from school were enrolled for comparison	1. 14-item questionnaire (12 related to voidings, 1 related to bowel habits, 1 related to stressful events, each score differently, with total score of 35)	1. Logistic regression for weighting of items 2. ROC curve analysis	1. No gender difference in DVISS total score 2. Children with 8.5 points or more were defined as cut-off point for dysfunctional voiding(90% sensitivity and 90% specificity)	1. Lacking of test-retest reliability , internal consistency 2. Different scores for different items, 3. Difficult to sum up total score 4. Binary and tertiary responses in different items
Assessment of urinary symptoms in children with	1. Age-matched case-controlled study 2. Group I(case,	Initial 35 items Exclude AUC<0.6 Finally, 15 items	1. ROC curve 2. Validation with extreme	1. Girls with DVSS of over 7.5 point (sensitivity	1. No test-retest reliability , internal consistency

dysfunctional elimination syndrome Tokgoz 2007[4]	n=81) children with dysfunctional voiding and constipation Group 2(non-case, n=102): aged matched healthy children without dysfunctional voiding and constipation	were included as final version	group	87.7%, specificity 94.2%) and boys with 4.5 points (93.8%, specificity 78%) selected as cut-off point for DVSS	2. Excluding items with Area under curve of less than 0.6 may hamper the content and construct validity of the questionnaire.
Development of a symptom score for dysfunctional elimination syndrome Afshar 2009[5]	1. Age- and gender-matched case-controlled study 2. Group 1(case, n=62) children with dysfunctional voiding and constipation 3. Group 2(non-case, n=50): healthy children without dysfunctional voiding and constipation randomly selected at ENT clinics	14 items, Likert scale Each item weighted equally (0-4 point for each question)	1. ROC 2.Discriminant function analysis 3.Cronbach's alpha for internal consistency 3.Pearson correlation for test-retest reliability 4.Factor analysis: 4 factors were noted 5.Question probing: do you think the item is easy	1. Cut-off point: 11(AUC0.903, 80% sensitivity and 91% Specificity) 2.Discriminant function analysis: 80% accuracy 3. Pearson correlation coefficient: 0.845 4.Four latent factors (incontinence, urgency, obstructive symptoms, and fecal soiling)	1. No comparison between genders 2. Using Pearson Correlation for test-retest reliability of total score without calculating reliability of individual items

## 2.2 問卷翻譯

問卷翻譯必須重視原有問卷獨創性之專屬權，所以做文獻回顧時，若發現已經有現存其他學者已經發展好的問卷，而且已經做過信度與效度的評估，應該可以考慮翻譯既有之問卷。翻譯問卷最主要的目的在於驗證翻譯版的問卷和原有的問卷是否有相同的效能 (equivalence)，其層面包含了 1.觀念上的等效 (conceptual equivalence)：在不同文化之下，同樣的問題或觀念是否相同，比如兒童膀胱逼尿肌過動造成急尿和憋尿造成膀胱過漲而造成尿急，其觀念及病因並不太相同。2.項目等效 (item equivalence)：主要是針對特定項目是否可以適合用在特定族群，比如在一個沒有現代廁所，只有茅坑的傳統社會，問他們小孩上完廁所會不會自己沖馬桶？並不適用在這個地區 3.語意上的等效 (semantic equivalence)：比如我們把”噓噓”當尿尿的意思，但若給其他使用中文的地區，他們可能不瞭解”噓噓”的意思。4.操作上的等效 (operational equivalence)：不同的問卷有不同的施行方式，比如有些為讓受試者自己填(self-administered)的，有些則是用面談(interview) 5.測量的等效 (measurement equivalence)：指的是翻譯的問卷是否和原本的問卷有相同的信度與效度，當然，這只有在問卷已經翻譯後才可以施行。[11]

Guillemin F et al (1993)建議做問卷翻譯時，最好能夠由兩個獨立的小組或人員將原有問卷翻譯成我們需要的語文 (forward translation)，負責翻譯的人需要清楚知道所有問卷項目的目的及意義，這樣可以讓翻譯者使用較多當地慣用的語言。經過專家會議，決定翻譯版問卷的內容之後，還需要再由另外兩組人員將翻譯成當地語言的問卷再翻譯成原文 (back translation)，這兩組人員並沒有看過原問卷而且並不知道這份問卷的目的為何。並將這兩份再翻譯的問卷寄給原問卷作者，請原作者評估是否與原問卷有不同之處，並詢問是否有需要修改之處。

不過由於翻譯問卷的人通常具有較好的教育程度，因此這些翻譯過的問卷通常要讓與你的目標族群 (target population) 教育程度相當的人來評估這份問卷是否有以下的問題：1. 問卷語言程

度太難：一般的要求為希望問卷的難度不要超過 12 歲的閱讀程度 2.模糊不清的語言 (ambiguity):比如問你有上廁所嗎?上廁所可以是尿尿,也可是解大便,容易令人混淆。3. 使用了專業用語(如醫學名詞)或行話 (jargon),讓一般人看不懂 4.帶有批判的話語 (value-laden words):比如你認為醫師賺太多錢嗎?5.問題過於冗長 (length of item): Holden et al (1985)發現當一個問題只有 10-20 字母時,他的效度是 70-80 字母問題的 4 倍以上。

### 2.3 問卷題目分數的比重及加權

我們此次採用的兒童排尿障礙問卷,共有十個題目,每題題目的分數都相同,為 0-3 分,個別症狀分數相同。雖然在計算上比較簡單,但是有些學者認為並非所有的item都一樣重要,有些項目及症狀對於診斷兒童排尿障礙會比其他症狀重要,因此這個比較重要的症狀,相對於其他症狀,應該佔比較大的分數比重。不過如果量測不同面向都使用不同的分數,那麼不同面向之間便無法比較。因此是否要對不同問題或項目加權 (Weight) 存在很多爭議。[11]

支持調整個別項目及症狀分數加權的學者認為可以使用迴歸係數來作個別症狀分數加權,如果依變項為連續變項,一般使用多變數迴歸,而本研究的依變項為二元變項 (Dichotomous parameters),即個案或非個案 (case or non-case),可以使用羅吉斯迴歸 (logistic regression),並依個別症狀的迴歸係數作為分析哪些症狀對於預測個案的效度較好,並可以用來作為調整權重的依據。

$$\text{Logit}(P_i) = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_K X_{Ki} + e_i \quad (2-1)$$

$P_i$ : 第  $i$  個個體為事件的機率

$X_{Ki}$ : 獨立項目選項

$\beta_0$ : 常數

$\beta_K$ : 第  $K$  個項目的迴歸係數值 (權重)

$e_i$ : 隨機誤差

不過如果每個問題都計算不同的權重，容易造成量表計算的繁複性。問題在於每個問題都給予不同權重時，真的一定可以增加量表的準確性嗎？其實是視情況而定。Wainer's (1976) 便建議把  $\widehat{\beta}_K$  值很接近 0 的項目去除掉即可，不一定要對每個項目加權。

另外一派學者則認為不需要對個別症狀加權，Lei 和 Skinner 等人便利用 Social Readjustment rating scale (SRRS) 來評估四種加權方式，對於問卷效度的影響[12]

1. 使用原始作者提出的加權方法。
2. 不加權，每個項目給予相同的分數。
3. 將原始作者提出的加權權重，隨機的用在其他項目上。
4. 給予每個項目亂數(0-100)進行加權。

最後他們發現，四個不同權重版本，最後加總的總分之間的相關係數高達 0.97。換言之，加權並不影響最終的總分。最後大家取得的共識為當一個量表的題目大於 40 個項目時，不需要做加權，小於 20-40 個項目時，加權可能會有一些效果。但如果  $\widehat{\beta}_K$ （第 K 個項目的迴歸係數值或權重）的大小都很接近的話，加權的效果也不大。

## 2.4 問卷信度的評估

當我們發展一份問卷量表時，我們希望這份量表對於我們想要估計的事物可已有準確的重複量測結果 (reproducibility)，因此需要檢驗問卷量表是否具有信度 (Reliability)。信度主要是反應測量誤差的大小，信度越高則誤差越小，反之則越大。亦可稱為 "objectivity"，

"reproducibility"，"stability"，"agreement"，"association"，"sensitivity" 和 "precision"。[11]

定義為：

$$\text{Reliability} = \frac{\text{Subject Variability}}{\text{Subject Variability} + \text{Measurement Error}} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2} \quad (2-2)$$

信度的範圍為 0 到 1 [除了 Cronbach's  $\alpha$  值可能為負值，假如量表為多向度 (multi-dimensional)，如果兩個向度為負相關，則  $\alpha$  值為負值]。

### 2.4.1 量表之內項目及問題的一致性

首先，我們注意到了量表內的一致性，當評估整個量表內項目的內部均質性及內部一致性 (homogeneity and internal consistency) 時，如果想要評估的量表其每個項目的選項大於兩個 (比如 DVSS 的每個症狀項目有 4 個選項，0 到 3)，或為連續性分佈時，可以使用 Cronbach's  $\alpha$  test [13]。

$$\text{Cronbach's } \alpha = \frac{K}{K-1} \left( 1 - \frac{\sum_{i=1}^K \sigma_i^2}{\sigma_T^2} \right) \quad (2-3)$$

$K$  = 問卷量表中有  $K$  個問題

$\sigma_i^2$  = 第  $i$  個問題分數的變異數

$\sigma_T^2$  = 所有分數的變異數

如果假設每個問題  $i$  分數的變異數相等，由 Cronbach's  $\alpha$  test 衍生出 Spearman-Brown prophecy

$$\alpha = \frac{K}{K-1} \left( 1 - \frac{\sum_{i=1}^K \sigma_i^2}{\sigma_T^2} \right) = \frac{Kr_{Avg}}{1 + (K-1)r_{Avg}} \quad (2-4)$$

$K$ ：問卷量表中問題個數。

$r_{Avg}$  = 問題之間的平均關係係數。

由 Spearman-Brown prophecy 公式可知，那麼如果問題項目個數 ( $K$ ) 越大時，則  $\alpha$  值會變大，而越接近 1。因此我們如果將信度只有 0.5 的量表項目或問題個數增加一倍，那麼信度會增加



到 0.67。

### 判斷 $\alpha$ 的標準

由於  $\alpha$  值受量表項目或問題的數目影響，一般認為如果可以接受的  $\alpha$  值為大於 0.7。Nunnally(1978)[14] 認為若此量表用來作為研究工具， $\alpha$  值應達 0.8，如果此量表用來作為臨床用途並且測量同一個面向時，則需達 0.9。Ponterotto 和 Ruckdeschel 等人[15] 則認為如果受試者約 100 人，而問題小於 7 項，那麼  $\alpha$  值 0.7 即可，而如果受試者超過 300 人，而問題大於 11 項，那麼  $\alpha$  值需要達到 0.9 即可。但如果  $\alpha$  值超過 0.9，也表示這些問題或項目太過一致，不需要這麼多的問題及項目。因此對於  $\alpha$  值標準之要求仍須視量表問題數目及樣本數大小調整。

### 2.4.2 組內相關係數 (intraclass correlation, ICC)

被量表評估的個體(informant)和評估個體的評估者(observer)之間的信度計算的計算方式，比如有三個評估者評估十個兒童的成績，所有的學生都要三個評估者評估，成績的分數由 1 至 10。成績如下表。利用重複測量 (repeated measurement) ANOVA 的方法來計算

下表為三個評估者給予十個兒童的分數

兒童	評估者 (甲)	評估者 (乙)	評估者 (丙)	平均
1	6	7	8	7
2	4	5	6	5
3	2	2	2	2
4	3	4	5	4
5	5	4	6	5
6	8	9	10	9
7	5	7	9	7
8	6	7	8	7
9	4	5	8	6
10	7	8	8	8
平均	5	6	7	6

以下為有關計算信度時各種統計量之計算：

評估者平方和 (平方和, Sum of square):

$$10 * [(6 - 6)^2 + (7 - 6)^2 + (5 - 6)^2] = 20$$

兒童平方和 (平方和, Sum of square):

$$3 * [(6 - 6)^2 + (5 - 6)^2 + \dots + (7 - 6)^2] = 114.0$$

誤差平方和 (平方和, Sum of square error):

$$(\text{Observed score} - \text{Expected score})^2 = 10.0$$

兒童之均方和 (Mean square) :

$$3 \sigma_{\text{children}}^2 + \sigma_{\text{error}}^2$$

評估者之均方和 (Mean square) :

$$10 \sigma_{\text{observer}}^2 + \sigma_{\text{error}}^2$$

均方誤 (Mean square of error):

$$\sigma_{\text{error}}^2$$

$\sigma_{\text{error}}^2$  (誤差變異):

$$10/(9 \times 2) = 0.56$$

$\sigma_{children}^2$  (兒童之變異):

$$(\text{評估者之均方和} - \text{均方誤})/3 = (114/9 - 0.56)/3 = 4.04$$

$\sigma_{observer}^2$  (評估者之變異):

$$(\text{兒童之均方和} - \text{均方誤})/10 = 0.94$$

信度 (Reliability, inter-rater reliability):

$$\frac{\sigma_{children}^2}{\sigma_{children}^2 + \sigma_{error}^2} = \frac{4.04}{4.04 + 0.56} = 0.88$$

Fisher (1925) 稱此種信度的計算方式為組內相關係數 (intraclass correlation, ICC) 的一種，與組間相關係數 (interclass correlation, 即 Pearson correlation coefficient) 不同。組內相關係數指的是同組之內，有重複測量的值產生，這些重複測量值之間的相關性，組間相關係數指的是不同組別之間的相關性。Shrout and Fleiss 等人進一步的將 ICC 分類為三種：第一種：評估每個受試者的觀察者或評估者為隨機選取，第二種：同一個觀察者或評估者評估所有受試者，而這個觀察者或評估者是從一個母群體中選取。第三種：同一個觀察者或評估者評估所有受試者，而只有一個觀察者或評估者。

另外一個重要的觀念為分辨一致性 (Consistency) 和絕對一致性 (Absolute agreement)。之前計算信度時，變異數只有考慮受試者的變異數，這便是一致性 (Consistency)，設定為固定評估者，亦即固定效應 (fixed effect)。但如果把評估者的變異數也考慮進去 (比如母群體有十個評估者，隨機選取來評估受試者)，那麼便是絕對的一致性 (Absolute agreement)，亦即隨機效應 (random effect)。

一致性 (Consistency) 的定義如下

$$\begin{array}{c}
 \text{class 2 ICC} \quad \text{One observer} \\
 \downarrow \quad \swarrow \\
 ICC2(C,1) = \frac{\sigma_{children}^2}{\sigma_{children}^2 + \sigma_{error}^2} \\
 \uparrow \\
 \text{consistency}
 \end{array}
 \tag{2-5}$$

絕對一致性 (Absolute agreement) 的定義如下

$$\begin{array}{c}
 \text{class 2 ICC} \quad \text{One observer} \\
 \downarrow \quad \swarrow \\
 ICC2(A,1) = \frac{\sigma_{children}^2}{\sigma_{children}^2 + \sigma_{observer}^2 + \sigma_{error}^2} \\
 \uparrow \\
 \text{Absolute agreement}
 \end{array}
 \tag{2-6}$$

多次觀察的一致性定義

$$ICC2(C,k) = \frac{k\sigma_{children}^2}{k\sigma_{children}^2 + \sigma_{error}^2} = \frac{MS_{children} - MS_{error}}{MS_{children}}
 \tag{2-7}$$

k=觀察的次數

多次觀察的絕對一致性定義

$$ICC2(A,k) = \frac{k\sigma_{children}^2}{k\sigma_{children}^2 + \sigma_{observers}^2 + \sigma_{error}^2} = \frac{MS_{children} - MS_{error}}{MS_{children} + (MS_{observers} - MS_{error}) / k}
 \tag{2-8}$$

k=觀察的次數

其他類型信度

ICC 可以用來評估觀察者或者評估者之間對於同樣刺激或者是受試者表現的評估一致性及信度，即觀察者間信度 (Inter-observer reliability)，如果是同一個觀察者或者評估者過一段時間對於同樣刺激或者是受試者表現的評估一致性及信度 (Intra-observer reliability)。有時候並沒

有觀察者或者評估者，只又受試者自己填量表 (self-administered)，那便是重測信度 (Test-retest reliability)，重測的時間與第一次測驗的信度約為 2 至 14 天，如果重測信度低有三種可能性 1. 這個量表信度較差, 2. 這個量表量測的建構可能會隨時間變動，如心情或疼痛，一週後便差異很大。3. 這個量表為有反應性的 (reactive)，受試者經過第一次量測後，回去會思考並修改反應，因此有較差的重測信度。

### 信度的準則

Nunnally (1978) [14]認為信度的標準，如果是使用在臨床研究這個量表的信度 0.7 即可，但如果是用在臨床需要達到 0.9。另外需要探討的是由量表的信度影響到的錯誤分組 (misclassification)，比如當使用一個量表將個案分類為個案及非個案時，信度為 0 時，有一半的機會重複測量結果會不相同，但如果信度是 0.5 時，有 37% 重複測量結果會不相同，而信度為 0.8 時仍有 20% 重複測量結果會不相同。若信度是 0.95 時，才會只有 2.2% 重複測量結果不相同。因此，0.75 的信度對一個測量工具為基本要求。

由信度的定義，我們可以減少隨機誤差的變異數，一般可以對評估者做訓練，減少評估者之間的差異，也可以增加實際變異數 (true variance)，比如當所有的個案都集中在極端值時，可以增加極端值的選項，來增加實際變異數。如果受試者的組成較異質化，那麼信度 (reliability) 會增加，因此另外一個不合法的作法是將一個量表使用在較異質化受試者族群，量表的信度便會增加，但這樣作會減少量表區辨個案及非個案的能力。

## 2.5 量表的效度評估

除了量表的信度 (reliability) 之外，即我們可以反覆使用此量表測量某個我們有興趣的事物

(attribute)，測量結果可以重複相同的表現出來，我們尚且需要知道這個可以被反覆測量出的結果是否可以反映出我們有興趣量測的事物。[11]

首先我們先從信度 (reliability) 的定義開始：

$$Reliability = \frac{\sigma_{True}^2}{\sigma_{Observed}^2} = \frac{\sigma_{CI}^2 + \sigma_{SE}^2}{\sigma_{CI}^2 + \sigma_{SE}^2 + \sigma_{RE}^2} \quad (2-9)$$

$\sigma_{Observed}^2$  = 全部分數(score)的變異數

$\sigma_{True}^2$  = 真實分數 (true score) 的變異數

$\sigma_{CI}^2$  = 有興趣的事物 (construct of interest) 的變異數

$\sigma_{SE}^2$  = 系統性誤差 (systemic error) 的變異數

$\sigma_{RE}^2$  = 隨機誤差 (random error) 的變異數

而效度的定義則為

$$Validity = \frac{\sigma_{CI}^2}{\sigma_{Observed}^2} = \frac{\sigma_{CI}^2}{\sigma_{CI}^2 + \sigma_{SE}^2 + \sigma_{RE}^2} \quad (2-10)$$

因此，信度一定大於等於效度( $0 \leq \text{validity} \leq \text{reliability} \leq 1$ )，信度為效度的上限。

效度主要分為三類：1. 內容效度 (content validity)，2. 效標關聯效度 (criterion validity)，3. 建構效度 (construct validity)。至於效度驗證 (Validation) 則為一個驗證假說測試的過程 (a process of hypothesis testing)，驗證一個量表測量出來的結果是不是我們真正想要測量的事物或面向，並藉由這個量表量測出的分數可以推論(inference)這個人有多少這個面向的特質。在闡述量表效度時，我們不能說這個量表已經被證實具有效度了，只能說這個量表已經被證實當使用在測量某個族群的某個特質或構想具有效度。

表面上的效度 (Face validity)：其定義為是否這些問題或項目是否在看起來就是要量測我們想要有興趣的事物。

內容效度 (Content validity，亦可稱為 content coverage 或 content relevance)：我們的問卷內容

是否能夠充分的包含研究題目的所有面向(domain)，尤其當我們遺漏了某個重要的組成，那我們對於此量表量測出的結果做出的推論便會有錯誤。

效標關聯效度 (Criterion validity)：一般來說即是這個量表可以多正確的預測他所要量測的事物。主要可以分成 1. 一致性效度 (concurrent validation) 2. 預測效度 (predictive validation)。

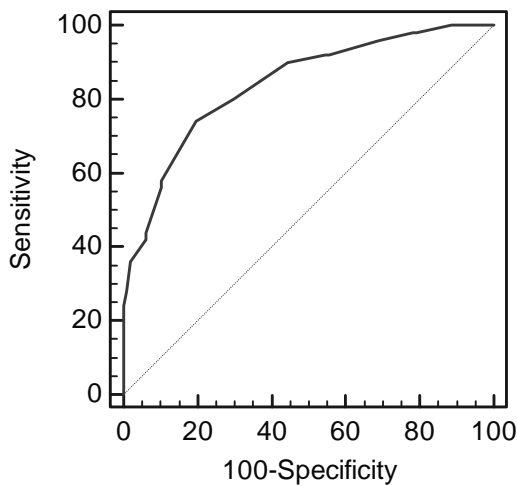
一致性效度：我們研究一個新量表和一個已被接受的舊量表相關性，便是一致性效度。

預測性效度：受試者經過黃金診斷標準或者確切的診斷標準分類為個案及非個案，我們評估利用此量表是否可以事先預測受試者為個案或非個案，即為此量表的預測性效度。

至於為什麼已經有一個很好的舊量表或診斷標準的情況下，還需要一個新的量表或診斷標準呢？最主要的原因為舊量表或診斷標準可能因為題目過多太花費時間，或者是昂貴，具有侵入性，對受試者有危險性或者是結果需要很長的時間才會知道，我們希望可以早一點預測結果，因此需要一個較簡易的量表或診斷標準。

利用 receiver operating characteristic curves (ROC curves, 如 Figure 2.5-1) 來找出較適切的切點來分辨受試者是不是具有這項疾病或者特質。當我們提高切點的分數時，可以減少偵測到非個案病人，但也有較多個案被遺漏，降低切點的分數時，可以偵測到較多個案病人，但也有較多非個案被偵測到。ROC 的縱軸為敏感性 (sensitivity) 而橫軸為 1-特異性 (1-specificity) 或者偽陽性率 (false positive rate)。而一個診斷方法的好壞則是視 ROC curves (Figure 2.5-1) 的 Area Under Curve (AUC) 的大小來決定診斷方法的好壞。

Figure 2.5-1 Receiver Operating Characteristic curve (ROC curve)



### 建構效度 (Construct validity)

比如當我們要測量焦慮、智商、疼痛的程度時，我們無法直接測量這些潛在變項 (latent variable)，但根據我們對於焦慮的假設，焦慮會有冒汗，心悸，來回踱步 (躁動)，失眠等等的表現，而我們可以觀察到這些表現，為可觀察變項 (observed variables)。由於這些可觀察變項為焦慮造成的結果，因此又稱為 effect indicator。這些潛在變項即為假設性的建構(hypothetical construct)，需要有他的理論 (theory) 去連結可觀察變項與建構之間的相關性。

### 以極端族群 (Extreme groups) 檢驗量表的效度

最簡單可以檢驗量表效度的方法為將量表量測兩組受試者，一組為個案 (具有我們有興趣以此量表量測的特質或行為) 一組為非個案，因此，個案應該在此量表會有較高的分數。這種方法又稱為以極端族群驗證建構效度 (construct validation by extreme groups) 又稱為區辨效度 (Discriminative validation)。

## 2.6 探索性因素分析 (exploratory factor analysis)

由於兒童排尿障礙患者的相關症狀呈相關性多變量分佈，因此因素分析 (factor analysis) 除了可以用來探討這些症狀 (觀察變項, observable variable) 的相關性，更可以依潛在變項 (latent



variables)將這些症狀進行分類。通常因素分析可以使用下列兩種方式施行A. 以一種探索性，且不帶有偏見的方式去檢驗觀察變項與潛在變項之間的相關性。B. 當作一個檢驗一個因素結構 (factor structure)，檢驗特定觀察變項與特定潛在變項間的相關性。[16]

因素分析 (Factor analysis) 的表示:

$$\begin{aligned} x_1 &= \lambda_{11}f_1 + \lambda_{12}f_2 + \dots + \lambda_{1k}f_k + u_1 \\ x_2 &= \lambda_{21}f_1 + \lambda_{22}f_2 + \dots + \lambda_{2k}f_k + u_2 \\ &\vdots \\ x_p &= \lambda_{p1}f_1 + \lambda_{p2}f_2 + \dots + \lambda_{pk}f_k + u_p \end{aligned} \tag{2-11}$$

觀察變項(observable variable)  $X' = [x_1, x_2, \dots, x_p]$

潛在變項(latent variables)  $F' = [f_1, f_2, \dots, f_k] (k < p)$

以向量表示為

$$x = \Lambda f + u = (\Lambda M)(M' f) + u \tag{2-12}$$

$\Lambda$  : 迴歸係數，又稱為因素負荷量(factor loadings)

$u_p$  : 不可被潛在變項解釋的變異， $u_1, \dots, u_p$  之間獨立不相關。

$M$  : 為  $k \times k$  因素的正交矩陣，因此  $(M \times M')$  為  $I$  (單位矩陣)

$f$ :  $k$  因素矩陣

由於潛在變項(latent variables)，我們無法直接觀察到，因此我們需要標準化其分佈，使其平均為 0 而變異數為 1。

而觀察變數  $x_i$  的變異數為

$$\sigma_i^2 = \sum_{j=1}^k \lambda_{ij}^2 + \psi_i \tag{2-13}$$

$\sigma_i^2$ : 觀察變數  $x_i$  的變異數

$h_i^2 = \sum_{j=1}^k \lambda_{ij}^2$  為共同性(communality)，指的是  $x_i$  藉由這些共同因子和其他因素共同解釋的變異數

$\psi_i$  為特異變異數 (specific variance)，指的是  $x_i$  無法藉由這些共同因子所解釋的變異數，因此  $\psi_i$  之間獨立互不相關。

$x_i$  和  $x_j$  之間的共變異數  $\sigma_{ij}$  為

$$\sigma_{ij} = \sum_{l=1}^k \lambda_{il} \lambda_{jl} \quad (2-14)$$

而  $x_i$  和  $x_j$  之間的共變異數矩陣 (covariance structure,  $\Sigma$ ) 為

$$\Sigma = \Lambda \Lambda' + \Psi = (\Lambda M)(\Lambda M)' + \Psi \quad (2-15)$$

$\Lambda$ ：迴歸係數，又稱為因素負荷量(factor loadings)

$\Psi$ ：為對角線以  $\psi_i$  組成的對角矩陣(diagonal matrix)。

$M$ ：為  $k \times k$  因素的正交矩陣 (orthogonal matrix)，因此  $(M \times M')$  為  $I$  (單位矩陣)

由式子 2-12 和 2-15 可知，只要  $M$  為正交矩陣， $x_i$  和  $x_j$  的共變異數矩陣  $\Sigma$  都可以被因素  $f$ ，其因素負荷量為  $\Lambda$ ，和新的因素  $f^* = M' f$  其因素負荷量為  $(\Lambda M)$  同等解釋，所以為了讓其因素的解為特定解 (unique solution)，需要給予限定條件 (constraints)，即為限定  $G = \Lambda' \Psi^{-1} \Lambda$  為一個對角線矩陣(diagonal matrix)，而且其對角線的元素排列為由大到小，類似主成分分析中的特徵值 (eigenvalue)，第一個因素 (factor) 對於原有變數所貢獻的變異數最多，依序次之，而且這些因素有互相獨立的性質(orthogonal properties)。

主因素分析 (Principal factor analysis)

主因素分析與主成分分析 (principal component analysis) 的技巧類似，只是主因素分析所採用的共變異數矩陣的對角線減去了特異變異數：

$$S^* = S - \hat{\Psi} \quad (2-16)$$

$S$  為原觀察變數  $x_i$  的共變異數矩陣 (covariance matrix)

$\hat{\Psi}$  為包含  $\psi_i$  的對角矩陣 (diagonal matrix)

$S^*$  為 reduced covariance，因此  $S^*$  其對角線的值則為估計的共同性，

我們最常用來運算共同性(communalities)的方法為

- a.  $X_i$  與其他觀察變數的 multiple correlation coefficients 的平方
- b.  $X_i$  與其他觀察變數相關係數 (correlation coefficients) 的極大值

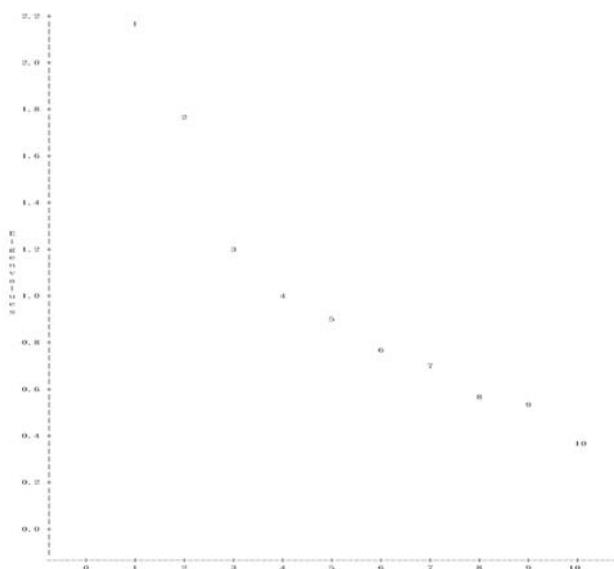
如果  $X_i$  與其他變數的相關性較大，則會有較大的共同性值。而主因素分析可能遇到的困難在有時候於計算出來的共同性值(communalities)可能會超出  $X_i$  的變異數，使得 specific variance( $\psi_i$ ) 的估計值可能為負值，而我們無法接受 specific variance 為負值，因此需要其他方法去調整。

通常我們決定要選取  $P^*$  個因素時，會考慮這  $P^*$  個因素可以解釋共同性(communalities)的比例 (proportion)。若要考慮第 1 到  $P^*$  個因素解釋了多少共同性值(communalities)的比例(cumulative proportion)，則可以下列式子表示：

$$P^* = \frac{\sum_{k=1}^{p^*} \lambda_k}{\text{trace}(S)} \quad (2-17)$$

另外 Cattell (1965) 提出了 scree diagram (Figure 2.6-1) 的觀念，依每個因素的 eigenvalue 由大到小作排列，選取  $P^*$  個因素至 scree diagram 有出現明顯轉折 (elbow) 為止。

Figure 2.6-1 scree diagram



### 因素旋轉(Rotation of factors)

由於因素之間互相獨立 (orthogonal property)，因此有可能造成下列兩種現象

1. 變數可能會在一個以上的因素有顯著的因素負荷 (factor loadings)，增加因素的複雜度 (factorial complexity)
2. 除了第一個因素之外，其他的因素的因素負荷會有正有負，即為雙極性(bipolar)

以上現象會造成解讀因素時的困難，最好是一個變數只在一個因素有較大的因素負荷，而在其他的因素最好是接近零，這便是 Thurstone (1931)提出的 simple structure 原則：

1. 因素負荷矩陣的每一列應該至少有一個負荷接近零。
2. 因素負荷矩陣的每一行至少應該有 k 個負荷接近零。
3. 每對因素負荷矩陣的行 (column) 至少有好幾個變數的負荷在一行接近零另一行數值較大。
4. 如果因素大於四個，則每一對的行 (column) 至少有好幾個變數的負荷都接近零。
5. 因素負荷矩陣中，每一對的行 (column) 只有少數幾個變數的負荷都不為零。

為了盡量符合 simple structure 的原則，可以旋轉因素的軸線 (axis)。最常使用的方法為 varimax rotation (Kaiser 1958)，經過 varimax rotation 之後的軸線為垂直的性質，其原則為讓較少的因

素有較大的因素負荷，而有較多的因素有接近零（near-zero）的因素負荷。

另一種為傾斜旋轉 oblique rotation，不過這種技巧產生的因素之間可能會有相關性。雖然我們可以隨著我們的需要選擇軸線的旋轉，如垂直或傾斜，但軸線（axis）旋轉並不會改變原有的變項的分佈，而且軸線的旋轉依然是從原點經過，我們只是讓這些點的分佈可以以較簡單的方式描述及解釋。

### 因素分數（Factor Scores）

經過因素分析後，有時候我們需要計算每一個個體觀察值在因素分析後個別因素的分數

因此在假設因素為常態分佈之下，因素的分佈以下面式子表示：

$$N\left[\Lambda' \Sigma^{-1}(x - \mu), (\Lambda' \Psi^{-1} \Lambda + I)^{-1}\right] \quad (2-18)$$

$\Lambda$ : 因素負荷矩陣

$\Sigma$  : 為原始變數共變異數矩陣（covariance）

$x$ : 個案原始分數矩陣

$\mu$  : 個案原始平均分數矩陣

$\Psi$ : 為包含特異變異的對角矩陣



因此利用下列式子計算每個個案在每個因素的分數：

$$f = \Lambda' S^{-1}(X - \bar{X}) \quad (2-19)$$

$f$ : 每個個案其於每個因素的分數矩陣

$\Lambda$ : 因素負荷矩陣

$S$  : 為原始變數的共變異數矩陣（covariance）

$X$ : 個案原始分數矩陣

$\bar{X}$  : 個案原始平均分數矩陣

## Chapter 3 研究方法

### 3.1 研究設計

本研究為一個年齡及性別相同配對的病例對照研究法 (age- and gender- matched case-controlled study)。有關病例組與對照組的定義如下：1. 病例組 (排尿障礙的兒童)：收集從西元 2008 年九月至 2011 年二月間，因為尿路感染，日間尿失禁或夜間尿床來某區域醫院小兒泌尿科門診求診的病童，經臨床評估及尿路動力學檢查結果評估為排尿障礙的兒童。我們從病歷收集其相關基本健康資料，其中包括了年齡，身高，體重，DVSS 問卷。此 DVSS 問卷填寫的方式為由主要照顧兒童的父母與孩童一起填寫，如果有問題的話可以向研究護士詢問相關問題。結果有 60 位病童納入研究。2. 對照組：從托兒所及國小收集與個案年齡及性別相同的健康兒童的兒童排尿障礙問卷及相關基本健康資料，問卷一樣由主要照顧兒童的父母與孩童一起填寫，並收集其尿流速及殘尿量檢查結果。3. 排除條件為：A. 病例組：具有先天性神經系統異常，如脊髓膜膨出 (meningomyelocele) 及先天性尿道解剖結構異常的兒童予以排除分析。B. 對照組：排除具有先天性神經系統疾病，尿路感染病史，相關膀胱及泌尿道異常病史，與尿流速圖形異常且殘尿量過高 (>20ml) 的兒童。

### 3.2 測量問卷翻譯及內容效度

於 2006 年已取得美國泌尿科醫學會同意，將 DVSS (Farhat 等人[2]) 中文化，並研究其信度與效度，DVSS 詳細內容參考 Farhat 等人之文章。中文版為楊緒棟等人所翻譯，主要的研究對象為排尿訓練後，有排尿障礙及下尿路症狀的兒童。另外會請相同年齡而沒有排尿障礙的兒童填寫問卷作為對照。請兩位專家將我們的中文問卷轉譯為英文 (back-translation)，一位為泌尿

科專家，另一位非泌尿科專家，並將這兩份英文問卷寄回給原作者，取得原作者的同意。此外，為了能夠維持問卷內容效度是否符合目前專家認定範疇，我們也將中文問卷寄予泌尿科醫學會小兒泌尿委員，詢問原問卷與中文問卷是否在每一個項目與原問卷在概念及語意上是否有所差異。並詢問每一個項目與有排尿障礙及白天尿失禁的兒童是否相關性，並且考慮是否會有語意不清或者不容易閱讀的項目。共有五個委員回覆問卷，大致上委員都同意中文問卷在概念及語意上與原問卷相同。關於第五題，原問卷為提到”每天只上廁所一至兩次”，被其他論文評論語意不清，所以我們改為”上廁所尿尿一至兩次”，以符合國際小兒尿失禁學會對於過少排尿頻率(小於等於每天兩次)的定義[17]



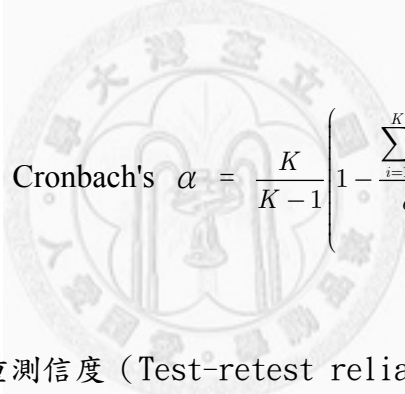
### 3.3 問卷信度分析

具有排尿障礙的小孩，於門診填寫 DVSS 後，於回家一星期中填寫了第二次 DVSS，在一星期後回診時繳回問卷，以利於採用組內相關係數(ICC)計算個案填寫 DVSS 之重測信度。

### 3.4 統計模型建立

#### 3.4.1 信度測試

根據第二章對於問卷信度評估中有關量表之內項目及問題之一致性，可以 Cronbach's alpha test (式子 2-3) 計算


$$\text{Cronbach's } \alpha = \frac{K}{K-1} \left( 1 - \frac{\sum_{i=1}^K \sigma_i^2}{\sigma_T^2} \right) \quad (3-1)$$

運用 2-5 式 ICC 運算，來計算重測信度 (Test-retest reliability):

$$ICC2(C,1) = \frac{\sigma_{children}^2}{\sigma_{children}^2 + \sigma_{error}^2} \quad (3-2)$$

C=consistency, 一致性

$\sigma_{children}^2$  = 受試者的變異數

$\sigma_{error}^2$  = 隨機誤差的變異數

Test-retest reliability = Inter-test reliability on the same subject = intra-children

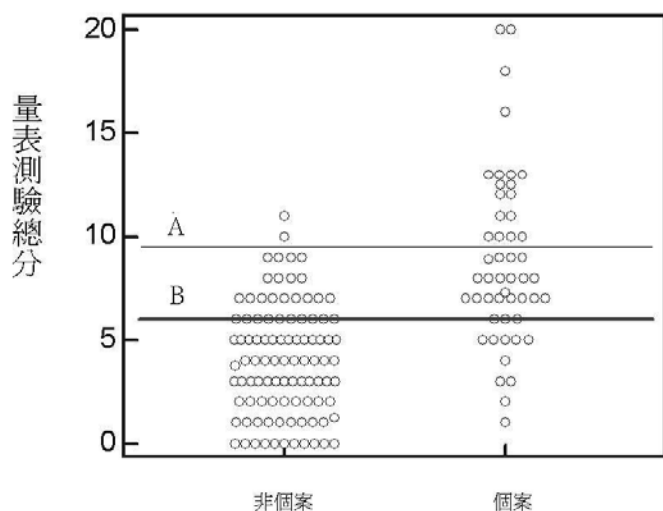


reliability (同一個兒童其重複測量 DVSS 結果的信度)

### 3.4.2 效度建立：

預測性效度：利用此量表可預測受試者（已經依確切的診斷結果分類為個案和非個案）是否為個案。利用 Receiver Operating Characteristic Curve Analysis 來比較不同模型的 Area Under the Curve (AUC)大小，並選擇不同的切點 (cut-off point) 來檢驗模型分辨個案及非個案的敏感性及特異性的值，如下圖(Figure 3.4-1)，當我們切點由 B 提高至 A，則會減少偽陽性及敏感度，但是會提高特異度及偽陰性。

Figure 3.4-1 選擇不同的切點 (cut-off point) 來分辨個案及非個案



為求預測效度可利用羅吉斯迴歸，來檢驗不同模型分辨個案及對照的模型適切度，模式表示如下：

$$\text{Logit}(P_i) = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_K X_{Ki} + e_i \quad (3-3)$$

$P_i$ ：第 i 個 subject 為個案的機率

$X_{ik}$ ：獨立項目選項

$\beta_0$ ：常數

$\beta_K$ ：第 K 個項目的迴歸係數值，即為風險比

$e_i$ ：隨機誤差



模型適切度及 ROC 的 AUC 值比較以 Akaike' s information criterion (AIC)=-2logL+2p 作為模型適切度 (model fitness) 的選擇，AIC 值越小，模型適切度越好。

logL=log(likelihood), 模型概似比取 log 值

p=模型所使用的參數數目

不同模型其 ROC curve 之 AUC 值比較，採用 DeLong 等人提出的無母數方法 [18]，表示方法如下：

$$(\hat{\theta} - \theta)L' \left[ L \left( \frac{1}{m} S_{10} + \frac{1}{n} S_{01} \right) L' \right]^{-1} L(\hat{\theta} - \theta) \quad (3-4)$$

$\theta$ : 當我們隨機從個案組中選出 i 個體，其分數 (DVSS) 大於隨機從對照組中選出 j 個體之分數 (DVSS) 的機率

$\hat{\theta}$ : 當我們隨機從個案組中選出 i 個體，其分數 (有興趣與 DVSS 比較之模型分數) 大於隨機從對照組中選出 j 個體之分數 (有興趣與 DVSS 比較之模型分數) 的機率

L: 係數的列向量 (row vector of coefficients)

$S_{10}$ : 個案 i 與對照組所有個體的共變異數 (covariance)

$S_{01}$ : 對照組 j 與所有個案組個體的共變異數 (covariance)

m: 個案組的數量

n: 對照組的數量

統計量分布為卡方分佈 (chi-square distribution)，而自由度為  $LSL'$  矩陣的秩 (rank)

### 3.4.3 因素分析 (factor analysis)

因素分析計算方法如同式子 2-11 至 2-16

$$\begin{aligned}
 x_1 &= \lambda_{11}f_1 + \lambda_{12}f_2 + \dots + \lambda_{1k}f_k + u_1 \\
 x_2 &= \lambda_{21}f_1 + \lambda_{22}f_2 + \dots + \lambda_{2k}f_k + u_2 \\
 &\dots \dots \dots \\
 x_p &= \lambda_{p1}f_1 + \lambda_{p2}f_2 + \dots + \lambda_{pk}f_k + u_p
 \end{aligned}
 \tag{3-5}$$

以矩陣  $X = \Lambda f + u$  表示

每個變數的變異數  $\sigma_i^2$ ，可以拆解成  $\sum_{j=1}^k \lambda_{ij}^2$  (共同性, communality) 和  $\psi_i$  (特異變異, specific variance)

$$\sigma_i^2 = \sum_{j=1}^k \lambda_{ij}^2 + \psi_i
 \tag{3-6}$$

使用主因素分析 (Principal factor analysis)，方法如同主成分分析 (principal component analysis)，但主因素分析利用 reduced covariance structure  $S^*$  計算

$$S^* = S - \hat{\Psi}
 \tag{3-7}$$

運算共同性(communalities)的方法為計算  $X_i$  與其他觀察變數的 multiple correlation coefficients 的平方計算。

因素的選取標準為要考慮第 1 到  $P^*$  個因素的特徵值解釋了多少共同性 (communalities) 的累積比例 (cumulative proportion)，計算如式子 2-17

$$P^* = \frac{\sum_{k=1}^{P^*} \lambda_k}{\text{trace}(S)} \quad (3-8)$$

找出因素後，在倒回來依標準化迴歸係數 (standardized regression coefficient) 計算每個個案及對照組兒童新的因素分數 (factor score)：(計算方法同式子 2-19)

$$f = \Lambda' S^{-1} (X - \bar{X}) \quad (3-9)$$

並以這些潛在因子去作羅吉斯迴歸及 ROC curve 分析，比較這些潛在變項對於分辨個案及非個案的模型適切度及預測效度。



#### 3.4.4 運用貝氏方法來調整 DVSS 的預測機率。

由於不同盛行率地區，利用 DVSS 來預測受試者是否患有排尿障礙 (DV) 需要依盛行率來調整其預測機率，因此將貝氏理論運用於調整 DVSS 的預測機率。

$$P(DV^+|DVSS) = \frac{P(DVSS|DV^+)P(DV^+)}{P(DVSS)} = \frac{P(DVSS|DV^+)P(DV^+)}{P(DVSS|DV^+)P(DV^+) + P(DVSS|DV^-)P(DV^-)} \quad (3-10)$$

事前機率 (Prior probability) 為 DV 的盛行率 ( $P(DV^+)$ )，根據文獻探討，兒童排尿障礙在七歲小孩子的盛行率約為 2 到 15%，至於在小兒泌尿科門診的盛行率則為 30%。

概似比(Likelihood)

假設 DVSS 在個案及對照組的分佈為常態分布 (normal distribution)，則 DV+和 DV-在此條件之下其概似函數分別為

$$\text{Likelihood} = P(\text{DVSS}|\text{DV}^+) = \frac{1}{\sqrt{2\pi\sigma_{\text{DVSS,DV}^+}^2}} \exp\left(\frac{-(\text{DVSS} - \mu_{\text{DVSS,DV}^+})^2}{2\sigma_{\text{DVSS,DV}^+}^2}\right) \quad (3-11)$$

$$\text{Likelihood} = P(\text{DVSS}|\text{DV}^-) = \frac{1}{\sqrt{2\pi\sigma_{\text{DVSS,DV}^-}^2}} \exp\left(\frac{-(\text{DVSS} - \mu_{\text{DVSS,DV}^-})^2}{2\sigma_{\text{DVSS,DV}^-}^2}\right) \quad (3-12)$$

$\mu_{\text{DVSS,DV}^+}$ : 患有排尿障礙兒童其 DVSS 的平均值

$\sigma_{\text{DVSS,DV}^+}^2$ : 患有排尿障礙兒童其 DVSS 的變異數

$\mu_{\text{DVSS,DV}^-}$ : 對照組兒童其 DVSS 的平均值

$\sigma_{\text{DVSS,DV}^-}^2$ : 對照組兒童其 DVSS 的變異數

以 DVSS 篩檢一個族群小孩，經由對於排尿障礙的盛行率調整，預測在此 DVSS 分數下，有排尿障礙的機率：

$$P(\text{DV}^+|\text{DVSS}) \propto \frac{P(\text{DVSS}|\text{DV}^+) \times P(\text{DV}^+)}{P(\text{DVSS}|\text{DV}^-) \times P(\text{DV}^-) + P(\text{DVSS}|\text{DV}^+) \times P(\text{DV}^+)} \quad (3-13)$$

而風險比 (odds) 為

$$\text{Posterior odds ratio} = \frac{P(DV^+|DVSS)}{P(DV^-|DVSS)} = \frac{P(DVSS|DV^+) \times P(DV^+)}{P(DVSS|DV^-) \times P(DV^-)} \quad (3-14)$$

論文所使用統計分析之軟體為 SAS® version 9.2



## Chapter 4 結果

### 4.1 基本資料描述性統計

病例組兒童的平均年齡為  $7.5 \pm 2.8$  歲，其中有 38 位女孩及 22 位男孩，而對照組的兒童的平均年齡為  $7.5 \pm 2.7$  歲，其中有 145 位女孩及 90 位男孩。如 Table 4.1-1。在個案組與對照組的十個題目及總分中，除了低排尿次數 (Q5) 之外，兩組的症狀包含了大便，小便及壓力事件的分數都有統計上顯著的差異。另外，個案組在尿流速檢查中有較低的最大尿流速及較多的殘尿量。根據過去的研究，有嚴重的排尿障礙的兒童常常會避免喝水，減少排尿及漏尿的次數，在考慮專家內容效度之下，合理的需要保留低排尿次數這個項目。而且當病童若有此症狀時，便可鼓勵他們多喝水，定時排尿，減少憋尿和尿路感染的可能性。依性別分層，單獨分析本研究中的女孩，個案與對照組之間僅 Q5(低排尿次數)沒有統計上顯著差異 (Table 4.1-2)。根據本研究中男孩單獨分析，由於個案較少，個案與對照組之間僅 Q1 (尿失禁)，Q2 (尿失禁量)，Q6 (急迫屈膝)，Q7(急尿感)及 Q8 (排尿困難)有統計上顯著差異 (Table 4.1-3)。



Table 4.1-1 個案及對照組的基本資料分佈與比較

Parameters	對照組 (n=235)	個案組 (n=60)	P-value
年齡 (years old)	7.5±2.7	7.5±2.8	0.90
性別 (男/女)	90/145	22/38	0.93
身高(cm)	128.4±17.1	127.8±18.1	0.52
體重 (kg)	28.6±10.7	28.3±13.1	0.71
Q1 (Urine incontinence)	0.09±0.32	1.00±0.84	<0.01
Q2 (Wetting amount)	0.08±0.34	0.76±0.86	<0.01
Q3 (Low defecation frequency)	0.72±0.69	0.98±0.70	0.01
Q4 (Difficult defecation)	0.68±0.78	1.06±0.95	<0.01
Q5 (Infrequent voiding)	0.22±0.57	0.18±0.47	0.88
Q6 (Curtsyng)	0.64±0.99	1.40±1.09	<0.01
Q7(Urgency)	0.65±0.88	1.73±1.14	<0.01
Q8 (Difficulty voiding)	0.05±0.21	0.69±0.96	<0.01
Q9 (Dysuria)	0.08±0.27	0.35±0.60	<0.01
Q10 (Stressful events)	0.99±1.41	1.48±1.50	0.01
總分 (DVSS)	4.13±2.60	9.65±3.87	<0.01
最大尿流速 Q <sub>max</sub> (ml/sec)	19.5±8.3	16.6±8.3	0.02
排尿量 Voided Volume (ml)	153.4±102.5	145.2±87.6	0.57
殘尿量 Post-void volume (ml)	6.9±11.5	18.1±24.9	<0.01

Table 4.1-2 女童個案及女童對照組的基本資料分佈與比較

Parameters	對照組(n=183)	個案組 (n=38)	p-value
年齡 (years old)	7.2±2.6	7.1±2.7	0.84
身高(cm)	126.3±17.3	124.1±18.0	0.63
體重 (kg)	26.8±9.3	27.3±12.2	0.82
Q1 (Urine incontinence)	0.11±0.35	1.1±0.8	<0.01
Q2 (Wetting amount)	0.10±0.40	0.88±0.94	<0.01
Q3 (Low defecation frequency)	0.72±0.69	0.95±0.61	0.03
Q4 (Difficult defecation)	0.69±0.79	1.16±1.02	0.003
Q5 (Infrequent voiding)	0.21±0.57	0.24±0.49	0.37
Q6 (Curtsyng)	0.63±1.00	1.44±1.08	<0.01
Q7(Urgency)	0.70±0.93	1.71±1.16	<0.01
Q8 (Difficulty voiding)	0.07±0.25	0.57±0.82	<0.01
Q9 (Dysuria)	0.11±0.31	0.45±0.56	<0.01
Q10 (Stressful events)	1.03±1.43	1.61±1.50	0.03
總分 (DVSS)	4.26±2.76	10.20±4.03	<0.01
最大尿流速 Qmax (ml/sec)	19.93±8.45	16.97±9.00	0.06
排尿量 Voided Volume (ml)	162.8±103.2	143.2±93.8	0.29
殘尿量 Post-void volume (ml)	7.4±11.7	21.8±30.1	<0.01

Table 4.1-3 男童個案及男童對照組的基本資料分佈與比較

Parameters	對照組(n=90)	個案組(n=22)	p-value
年齡 (years old)	8.06±2.72	8.13±2.76	0.91
身高(cm)	131.6±16.4	129.9±16.7	0.75
體重 (kg)	31.2±12.0	29.0±12.3	0.52
Q1 (Urine incontinence)	0.07±0.25	0.68±0.78	<0.01
Q2 (Wetting amount)	0.06±0.23	0.55±0.67	<0.01
Q3 (Low defecation frequency)	0.74±0.70	1.04±0.84	0.12
Q4 (Difficult defecation)	0.65±0.77	0.87±0.78	0.14
Q5 (Infrequent voiding)	0.23±0.58	0.09±0.42	0.14
Q6 (Curtsyng)	0.66±0.98	1.33±1.12	0.01
Q7(Urgency)	0.57±0.78	1.77±1.15	<0.01
Q8 (Difficulty voiding)	0.016±0.117	0.90±1.15	<0.01
Q9 (Dysuria)	0.03±0.18	0.18±0.66	0.23
Q10 (Stressful events)	0.90±1.38	1.23±1.51	0.36
總分 (DVSS)	3.92±2.31	8.70±3.48	<0.01
最大尿流速 Qmax (ml/sec)	18.7±8.1	15.9±7.0	0.11
排尿量 Voided Volume (ml)	138.3±100.0	148.7±77.8	0.65
殘尿量 Post-void volume (ml)	6.1±11.3	11.7±8.5	0.02

## 4.2 遺漏值處理

在 60 位有排尿障礙的兒童的排尿障礙問卷中，共有 600 個項目，遺漏的項目為 16 欄（佔 2.66%），處理方法為根據同一個個案填寫其他項目的平均值，作為這遺漏項目的值。各個項目遺漏的次數如 Table 4.2-1 遺漏最多的題目為第六題（急尿屈膝，Curtsyng：因急尿而有蹲下來，避免尿濕褲子的姿勢），共有 4 個遺漏值。

Table 4.2-1 問題一至問題十的遺漏值數量

題目	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
遺漏數量	0	2	0	3	0	4	0	3	3	1



## 4.3 信度分析

### 4.3.1 重測信度 (Test-retest reliability)

依組內相關係數(ICC)計算 29 位填寫兩次問卷的兒童，其 DVSS 總分及個別項目之重測信度，結果列於 Table 4.3 1。除了第六題（急尿屈膝 Curtsyng）的重測信度較差（ICC=0.60）之外，其中大部分的重測信度都介於 0.70 至 0.87 之間。至於總分的重測信度則為 0.89。

Table 4.3-1 個別問卷項目及總分的重測信度（組內相關係數）

Question	ICC consistency
Q1 (Urine incontinence)	0.79
Q2 (Wetting amount)	0.71
Q3 (Low defecation frequency)	0.72
Q4 (Difficult defecation)	0.72
Q5 (Infrequent voiding)	0.78
Q6 (Curtsyng)	0.60
Q7 (Urgency)	0.72
Q8 (Difficult voiding)	0.85
Q9 (Dysuria)	0.87
Q10 (Stressful events)	0.70
Total score	0.89

### 4.3.2 內部一致性 (internal consistency)

如同 Table 4.3-2 顯示,如果以 Cronbach's alpha test 測試整份問卷的內部一致性 (internal consistency),則個案全部項目的 Alpha 值,為 0.448。可見排尿障礙的症狀組成異質性 (heterogeneity) 較高,就算將症狀拆解成小便相關及大便相關其 Alpha 值仍舊不高,分別為 0.452 及 0.431。同樣的在對照組的 Alpha 值一樣不高,我們接著會採用因素分析來分析這些資料,可以將這些項目予以分類。

Table 4.3-2 個案及對照組項目的內部一致性 (internal consistency) 之 Cronbach's alpha 值

Items included	<i>α</i> coefficient of case	<i>α</i> coefficient of control group
Total items	0.448	0.237
Voiding-related items Q1, Q2, Q5, Q6, Q7, Q8, Q9	0.452	0.335
Defecation-related items Q3, Q4	0.431	0.554

## 4.4 預測效度分析 (predictive validity)

### 4.4.1 切點選擇

以 DVSS 總分作 ROC curve 分析，評估 DVSS 總分在不同切點之下預測所有受試者為個案或非個案的效度，其 AUC 值為 0.89 (Figure 4.4-1)。Table 4.4-1 顯示不同切點有著不同的敏感性 (sensitivity) 及特異性 (specificity)。一般希望兼顧到敏感性及特異性，所以取 ROC curve 最左上角的點作為切點。我們選擇  $>6.66$  作為切點，此時可以有 81.67% (95% CI: 69.6 - 90.5) 的敏感性及 82.13% (95% CI: 76.6 - 86.8) 的特異性。依性別分層分析，女童的 ROC curve analysis 發現其 AUC 值為 0.89 (Figure 4.4-2)。如果選擇 DVSS  $>8$  作為切點，可以有 71.05% (95% CI: 54.1 - 84.6) 的敏感性及較高的 93.79% (70.3 - 84.4) 的特異性，但我們選擇  $>6.66$  作為切點，犧牲了特異性換取較好的敏感性，此時可以有 81.58% (95% CI: 65.7 - 92.3) 的敏感性及 77.93% (70.3 - 84.4) 的特異性 (Table 4.4-2)。至於男童，其 AUC 為 0.91 (Figure 4.4-3)。以  $>6.66$  作為切點，此時可以有較好的敏感性 (81.82%, 95% CI: 59.7 - 94.8) 及特異性 (88.89%, 95% CI: 80.5 - 94.5) (Table 4.4-3)。因此，本研究男女的切點選擇並沒有差別。

Figure 4.4-1 建立預測效度的 ROC 曲線圖形 (所有個案)

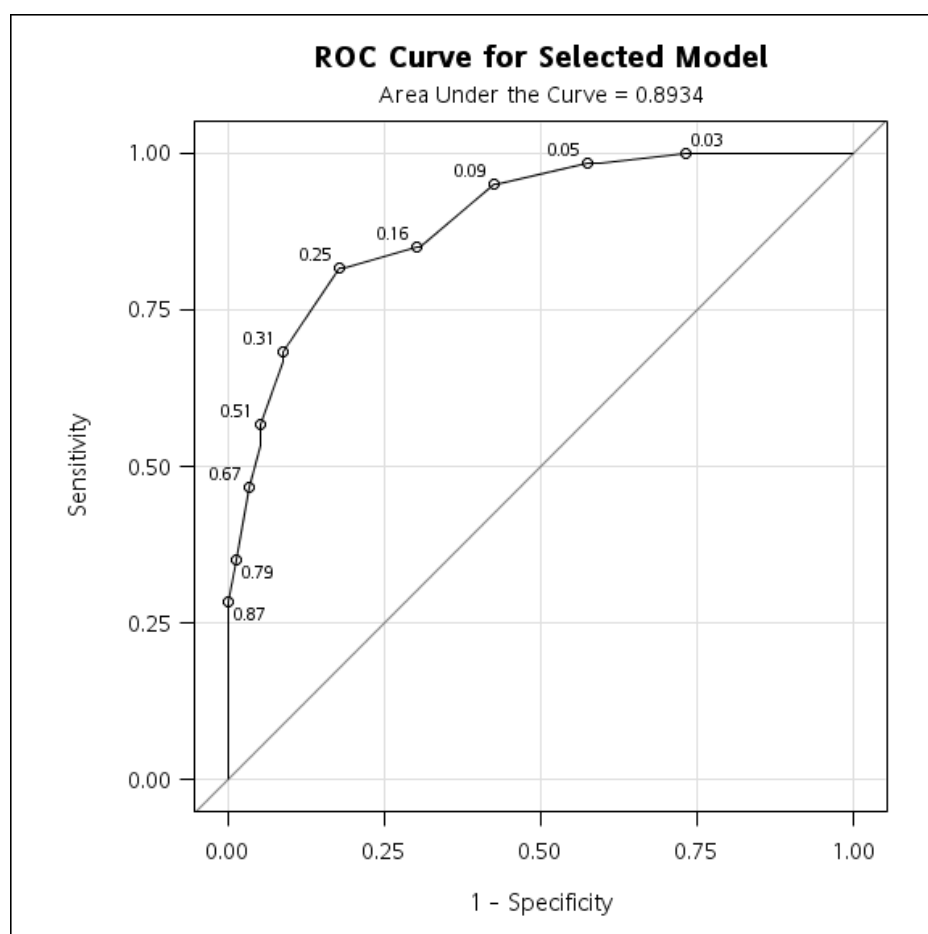


Table 4.4-1 選擇不同總分切點的敏感性及特異性 (所有個案)

Cut-off for DVSS	Sensitivity	95% confidence interval of sensitivity	Specificity	95% confidence interval of specificity
>=0	100.00	94.0 - 100.0	0.00	0.0 - 1.6
>4	95.00	86.1 - 99.0	57.02	50.4 - 63.4
>6	81.67	69.6 - 90.5	81.28	75.7 - 86.1
>6.66	81.67	69.6 - 90.5	82.13	76.6 - 86.8
>8	56.67	43.2 - 69.4	94.89	91.3 - 97.3
>10	35.00	23.1 - 48.4	98.72	96.3 - 99.7
>20	0.00	0.0 - 6.0	100.00	98.4 - 100.0



Figure 4.4-2 以羅吉斯迴歸建立預測效度的 ROC 曲線圖形 (女童)

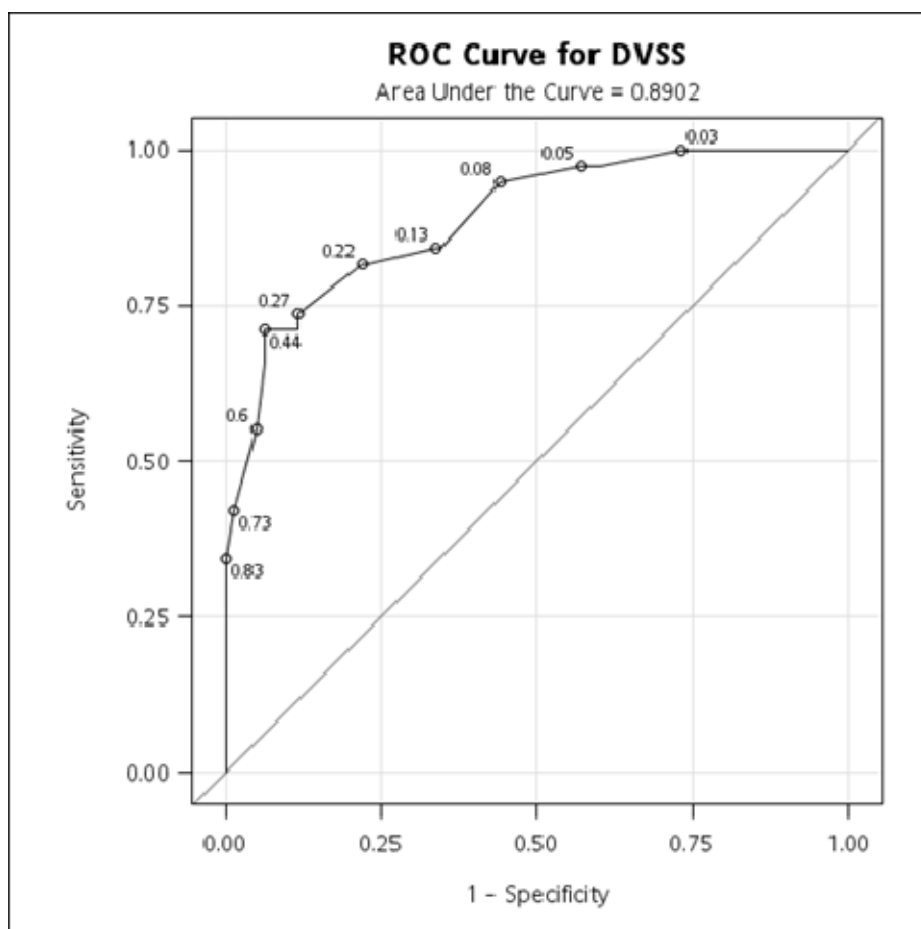


Table 4.4-2 選擇不同總分切點的敏感性及特異性 (女童)

Cut-off for DVSS	Sensitivity	95% confidence interval of sensitivity	Specificity	95% confidence interval of specificity
>4	94.74	82.3 - 99.4	55.86	47.4 - 64.1
>6	81.58	65.7 - 92.3	77.24	69.5 - 83.8
>6.66*	81.58	65.7 - 92.3	77.93	70.3 - 84.4
>8	71.05	54.1 - 84.6	93.79	88.5 - 97.1
>10	42.11	26.3 - 59.2	98.62	95.1 - 99.8
>20	0.00	0.0 - 9.3	100.00	97.5 - 100.0

Figure 4.4-3 以羅吉斯迴歸建立預測效度的 ROC 曲線圖形 (男童)

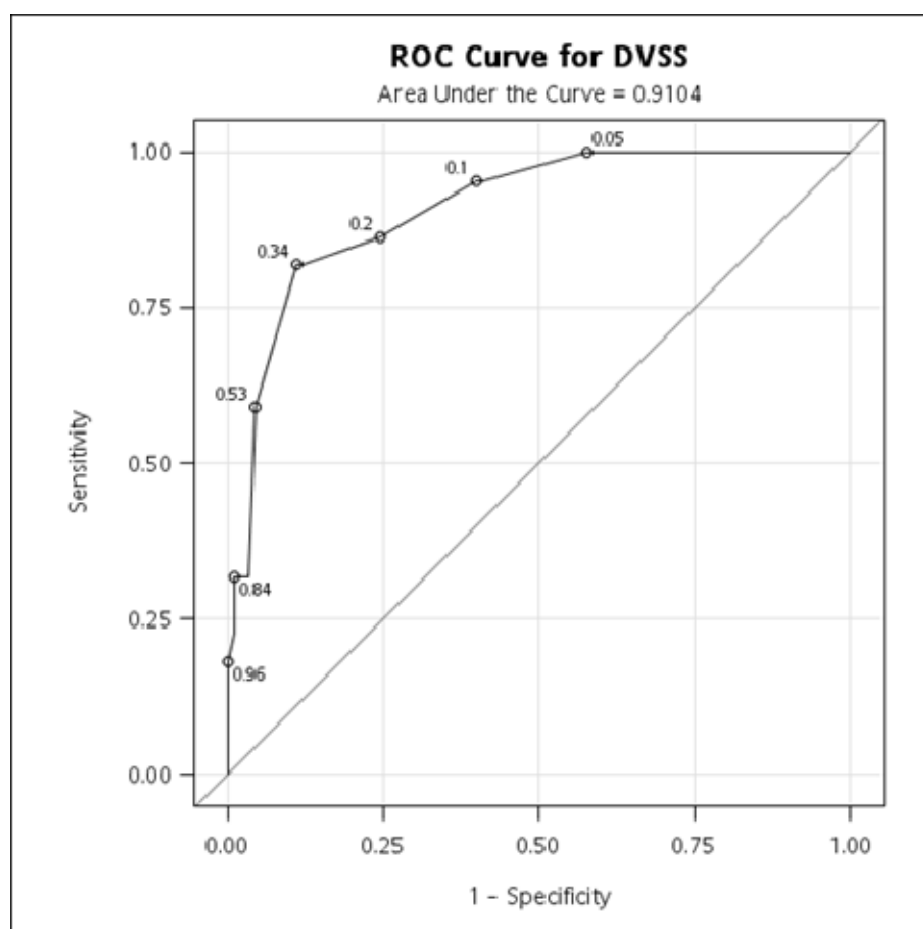


Table 4.4-3 選擇不同總分切點的敏感性及特異性 (男童)

Cut-off for DVSS	Sensitivity	95% confidence interval of sensitivity	Specificity	95% confidence interval of specificity
>4	95.45	77.2 - 99.9	58.89	48.0 - 69.2
>6	81.82	59.7 - 94.8	87.78	79.2 - 93.7
>6.66 *	81.82	59.7 - 94.8	88.89	80.5 - 94.5
>7	59.09	36.4 - 79.3	95.56	89.0 - 98.8
>8	31.82	13.9 - 54.9	96.67	90.6 - 99.3
>10	22.73	7.8 - 45.4	98.89	94.0 - 100.0
>20	0.00	0.0 - 15.4	100.00	96.0 - 100.0

## 4.5 主因素分析(principal factor analysis)

### 4.5.1 依個案症狀作主因素分析 (Principal factor analysis)

主因素分析為以 reduced covariance matrix 計算特徵值，以排尿障礙兒童的症狀作主因素分析，根據 Table 4-10，我們依據 proportion criterion (即抽取出的因素可以解釋共同性的比例)選取了三個因素，Table 4.5-2 為依個案症狀作因素分析的原始因素負荷，共同性及特異變異，經 varimax 旋轉 (Table 4.5-3)，依因素負荷將十個問題分類在 3 個因素：第一個因素 (Factor 1) :尿失禁及急尿症狀 (overactive bladder symptoms) Q1 尿失禁頻率,Q2 尿失禁量,Q6 屈膝急尿及 Q7 急尿症狀，第二個因素 (Factor 2) :排泄困難症狀 (difficult elimination symptoms)，Q3 低排便次數,Q4 排大便需用力,及 Q8 小便需用力，第三個因素 (Factor 3) :壓力及疼痛症狀(stress and dysuria symptoms)，Q5 排尿次數減少，Q9 解尿疼痛，及 Q10 壓力事件。

Table 4.5-1 依個案症狀作因素分析之因素特徵值 (個案組) 及其解釋共同性比例

因素	特徵值 ( Eigenvalue )	因素特徵值解釋共同性比例 (Proportion)	因素特徵值累積解釋共同性比例 (Cumulative proportion)
1	1.44	0.50	0.50
2	1.19	0.41	0.92
3	0.63	0.22	1.15
4	0.42	0.14	1.30
5	0.07	0.02	1.32
6	0.03	0.01	1.33
7	-0.14	-0.05	1.28
8	-0.19	-0.06	1.21
9	-0.29	-0.10	1.11
10	-0.32	-0.11	1.00

Table 4.5-2 個案症狀因素分析的原始因素負荷，共同性及特異變異（因素負荷>0.3, \*）

Items	Factor 1 (overactive bladder symptoms) loadings	Factor 2 (difficult elimination symptoms) loadings	Factor 3 (stress and dysuria symptoms) Loadings	Communality	Specific variance
Q1 (Urine incontinence)	0.68*	-0.25	0.13	0.54	0.46
Q2 (Wetting amount)	0.61*	-0.29	0.14	0.48	0.52
Q3 (Low defecation frequency)	-0.05	0.43*	0.27	0.26	0.74
Q4 (Difficult defecation)	0.31*	0.45*	0.21	0.34	0.66
Q5 (Infrequent voiding)	0.38*	0.12	-0.34*	0.28	0.72
Q6 (Curtsyng)	0.26	-0.23	0.18	0.15	0.85
Q7(Urgency)	0.27	-0.01	0.16	0.10	0.90
Q8 (Difficult voiding)	0.29	0.57*	0.05	0.42	0.58
Q9 (Dysuria)	0.31*	0.44*	-0.41*	0.46	0.54
Q10 (Stressful event)	0.20	-0.26	-0.36*	0.24	0.76
Total variance	1.44	1.10	0.63		

Table 4.5-3 個案症狀因素分析，經過 Varimax 軸線旋轉後的因素負荷，共同性及特異變異(因素負荷>0.3, \*)

Items	Factor 1 (overactive bladder symptoms) loadings	Factor 2 (difficult elimination symptoms) loadings	Factor 3 (stress and dysuria symptoms) Loadings	Communality	Specific variance
Q1 (Urine incontinence)	0.71 *	-0.06	0.18	0.54	0.46
Q2 (Wetting amount)	0.67 *	-0.09	0.11	0.48	0.52
Q3 (Low defecation frequency)	-0.10	0.50 *	-0.04	0.26	0.74
Q4 (Difficult defecation)	0.17	0.52 *	0.21	0.34	0.66
Q5 (Infrequent voiding)	0.13	-0.06	0.51 *	0.28	0.72
Q6 (Curtsyng)	0.38 *	-0.06	-0.08	0.15	0.85
Q7(Urgency)	0.30 *	0.11	0.02	0.10	0.90
Q8 (Difficult voiding)	0.04	0.53 *	0.37	0.42	0.58
Q9 (Dysuria)	-0.08	0.16	0.65 *	0.46	0.54
Q10 (Stressful event)	0.12	-0.40 *	0.26	0.24	0.76
Total variance	1.27	1.02	0.98		

#### 4.5.2 以對照組兒童的症狀作主因素分析 (Principal factor analysis)

以對照組兒童的症狀作主因素分析，由 Table 4.5-4 根據因素特徵值（對照組）及其解釋共同性比例，我們採用二個因素作因素分析的結果列於 Table 4.5-5，經過 varimax 旋轉（Table 4.5-6），依據因素負荷將十個問題分為 2 個因素，第一個因素（Factor 1）：尿失禁及急尿症狀（overactive bladder symptoms）包含了 Q1 尿失禁頻率，Q2 尿失禁量及 Q7 急尿症狀，主要是與膀胱過動及逼尿肌過動相關症狀，第二個因素（Factor 2）：排泄症狀（elimination symptoms）包含了 Q3 低排便次數，Q4 排大便需用力，Q5 排尿次數減少，Q8 小便需用力及 Q9 解尿疼痛。Q6 急尿屈膝和 Q10 壓力事件並不包含兩個因素之中

Table 4.5-4 以對照組兒童的症狀因素分析之因素特徵值（對照組）及其解釋共同性比例

因素	特徵值 ( Eigenvalue )	因素特徵值解釋共同性比例 (Proportion)	因素特徵值累積解釋共同性比例 (Cumulative proportion)
1	1.39	0.65	0.65
2	1.03	0.48	1.14
3	0.40	0.19	1.33
4	0.15	0.07	1.40
5	0.06	0.03	1.43
6	-0.051	-0.02	1.41
7	-0.10	-0.04	1.36
8	-0.21	-0.10	1.26
9	-0.25	-0.11	1.14
10	-0.30	-0.14	1.00

Table 4.5-5 對照組症狀因素分析的原始因素負荷，共同性及特異變異(因素負荷>0.3, \*)

Items	Factor 1 (overactive bladder symptoms) loadings	Factor 2 (elimination symptoms) Loadings	Communality	Specific variance
Q1 (Urine incontinence)	0.46*	-0.25	0.45	0.55
Q2 (Wetting amount)	0.46*	-0.29	0.48	0.52
Q3 (Low defecation frequency)	0.42*	0.43*	0.33	0.67
Q4 (Difficult defecation)	0.52*	0.45*	0.33	0.67
Q5 (Infrequent voiding)	0.36*	0.12	0.15	0.85
Q6 (Curtsyng)	0.25	-0.23	0.09	0.91
Q7(Urgency)	0.37*	-0.01	0.18	0.82
Q8 (Difficult voiding)	0.28	0.57*	0.20	0.80
Q9 (Dysuria)	0.19	0.44*	0.15	0.85
Q10 (Stressful event)	-0.28	-0.26	0.08	0.92
Total variance	1.39	1.04		

Table 4.5-6 對照組症狀因素分析，經過 Varimax 軸線旋轉後的因素負荷，共同性及特異變異（因素負荷>0.3, \*）

items	Factor 1 (overactive bladder symptoms) loadings	Factor 2 (elimination symptoms) loadings	Communality	Specific variance
Q1 (Urine incontinence)	0.67*	-0.05	0.45	0.55
Q2 (Wetting amount)	0.69*	-0.06	0.48	0.52
Q3 (Low defecation frequency)	0.04	0.58*	0.33	0.67
Q4 (Difficult defecation)	0.20	0.54*	0.33	0.671
Q5 (Infrequent voiding)	0.16	0.35*	0.15	0.85
Q6 (Curtsyng)	0.08	0.28	0.09	0.91
Q7(Urgency)	0.40*	0.12	0.18	0.82
Q8 (Difficult voiding)	-0.04	0.44*	0.20	0.80
Q9 (Dysuria)	0.09	0.37*	0.15	0.85
Q10 (Stressful event)	-0.24	-0.15	0.08	0.92
Total variance	1.23	1.20		



#### 4.5.3 以所有個案及對照組症狀作主因素分析 (Principal factor analysis)

將所有個案及對照組兒童的症狀作主因素分析，由 Table 4.5-7 根據因素特徵值及其解釋共同性比例，我們採用二個因素作因素分析的結果列於 Table 4.5-8，經過 varimax 旋轉 (Table 4.5-9)，依據因素負荷將十個問題分為 2 個因素，第一個因素 (Factor 1)：尿失禁及急尿症狀 (overactive bladder symptoms) 包含了 Q1 尿失禁頻率, Q2 尿失禁量, Q6 急尿屈膝，和 Q7 急尿症狀，主要是與膀胱過動及逼尿肌過動相關症狀，第二個因素 (Factor 2)：排泄症狀 (elimination symptoms) 包含了 Q3 低排便次數, Q4 排大便需用力, Q5 排尿次數減少，Q8 小便需用力及 Q9 解尿疼痛。其因素分析結果與對照組兒童的症狀作主因素分析類似，只有第一個因素多了 Q6 急尿屈膝。同樣的 Q10 壓力事件並不包含兩個因素之中。

Table 4.5-7 所有個案及對照組症狀因素分析之因素特徵值 (對照組) 及其解釋共同性比例

因素	特徵值 ( Eigenvalue )	因素特徵值解釋共同性比例 (Proportion)	因素特徵值累積解釋共同性比例 (Cumulative proportion)
1	2.12	0.78	0.78
2	0.83	0.30	1.08
3	0.39	0.14	1.23
4	0.14	0.05	1.28
5	0.03	0.01	1.29
6	-0.034	-0.01	1.28
7	-0.10	-0.04	1.24
8	-0.14	-0.05	1.19
9	-0.24	-0.08	1.10
10	-0.28	-0.10	1.00

Table 4.5-8 所有個案及對照組症狀作因素分析的原始因素負荷，共同性及特異變異(因素負荷 >0.3, \*)

Items	Factor 1 (overactive bladder symptoms) loadings	Factor 2 (elimination symptoms) Loadings	Communality	Specific variance
Q1 (Urine incontinence)	0.74 *	-0.35 *	0.68	0.32
Q2 (Wetting amount)	0.66 *	-0.37 *	0.57	0.43
Q3 (Low defecation frequency)	0.29	0.45 *	0.29	0.71
Q4 (Difficult defecation)	0.44 *	0.38 *	0.35	0.65
Q5 (Infrequent voiding)	0.17	0.28	0.11	0.89
Q6 (Curtsyng)	0.35 *	0.05	0.13	0.87
Q7(Urgency)	0.51 *	-0.10	0.27	0.73
Q8 (Difficult voiding)	0.53 *	0.22	0.34	0.66
Q9 (Dysuria)	0.37 *	0.19	0.18	0.82
Q10 (Stressful event)	0.00	-0.18	0.03	0.97
Total variance	2.12	0.83		

Table 4.5-9 所有個案及對照組症狀作因素分析，經過 Varimax 軸線旋轉後的因素負荷，共同性及特異變異（因素負荷>0.3, \*）

items	Factor 1 (overactive bladder symptoms) loadings	Factor 2 (elimination symptoms) loadings	Communality	Specific variance
Q1 (Urine incontinence)	0.82 *	0.01	0.67	0.33
Q2 (Wetting amount)	0.75 *	-0.04	0.57	0.43
Q3 (Low defecation frequency)	0.06	0.53 *	0.29	0.71
Q4 (Difficult defecation)	0.23	0.54 *	0.35	0.65
Q5 (Infrequent voiding)	0.02	0.33 *	0.11	0.89
Q6 (Curtsyng)	0.29	0.20	0.13	0.87
Q7(Urgency)	0.50 *	0.12	0.27	0.83
Q8 (Difficult voiding)	0.38 *	0.43 *	0.34	0.66
Q9 (Dysuria)	0.25	0.33 *	0.18	0.82
Q10 (Stressful event)	0.081	-0.16	0.03	0.87
Total variance	1.86	1.08		

## 4.6 羅吉斯迴歸模型之AUC值及模型適切度

經由羅吉斯迴歸將不同變項帶入，評估不同模型對於分辨個案與非個案的 ROC curve 其 AUC 值及評估模型適切度的 AIC 值 (Table 4.6-1)。模型一為原始模式，將年齡，性別及 DVSS 總分帶入模型 (AIC=182.8)，結果 AUC 值為 0.89 (95CI: 0.85-0.94)，女童發生排尿障礙的風險比 (和男生相比) 雖為 1.7 (95CI: 0.76-3.84)，但是因為收集個案數不夠，所以並沒有統計顯著差異。

模型二為將年齡，性別及 DVSS 個別症狀項目(Q1-Q10)帶入模型 (AIC=152.6)，AUC 值上升為 0.94 (95CI: 0.91-0.98)，十個症狀中以 Q1 (OR:10.1,95CI:3.41-30.19), Q6 (OR:1.7, 95CI:1.10-2.60), Q7(OR:2.2, 95CI:1.38-3.60), Q8 (OR:10.6, 95CI:3.46-32.67)和 Q10(OR:1.5, 95CI:1.08-2.15)風險比達統計顯著意義，而其中 Q1 白天尿失禁和 Q8 排尿困難風險比又特別的高。

模型三：如果從排尿障礙兒童的症狀作主因素分析，找出來的三個潛在因素 1.尿失禁及急尿症狀,2.排泄困難症狀，3.疼痛及壓力症狀和性別及年齡帶入模型 (AIC=195.3)，發現 AUC 值為 0.89 (95CI: 0.85-0.94)，其中因素一到因素三的風險比都有統計意義，但以因素一：尿失禁及急尿症狀風險比較高 (15.8, 95CI: 7.57-32.91)。

模型四：如果將對照組兒童的症狀作主因素分析找出來的二個潛在因素帶入模型 (AIC=170.7)，發現 AUC 值為 0.89 (95CI: 0.85-0.94)，其中因素一及因素二的風險比都有統計顯著意義

模型五：如果將所有個案組與對照組兒童的症狀作主因素分析找出來的二個潛在因素帶入

模型(AIC=168.4)，發現 AUC 值為 0.91 (95 CI: 0.85-0.94) 較高，其中因素一及因素二的風險比都有統計顯著意義

Table 4.6-2為比較各個模型其AUC值與模型一的AUC值差別，其統計量為卡方分佈。結果只有模型二與模型一有統計上的顯著差別，由於模型二帶入了十個症狀項目，對不同的症狀分數加權，所以增加了0.05的AUC值。而模型三至模型五與模型一的AUC值並沒有統計顯著差別，可見得由個案及對照組的症狀作因素分析後得到的因素，也可以很好的預測效度。



Table 4.6-1 各個模型 Receiver Operating Characteristic (ROC) curve 的 AUC (Area Under Curve) 值，模型適切度(Akaike' s information criterion)，變數迴歸係數及變數風險比之比較。

Parameters included in Model	AUC	AIC	Regression coefficient (standard error)	Odds ratio (95% confidence interval)					
Model 1: Age+gender+ DVSS	0.89	182.8	Age	0.09 (0.08)	Age	1.1 (0.94-1.27)			
			Gender	0.54 (0.41)	Gender	1.7 (0.76-3.84)			
			DVSS	0.65 (0.08)	DVSS	1.9 (1.60-2.26)			
			Age	0.13 (0.10)	Age	1.1 (0.93-1.38)			
			Gender	0.68 (0.53)	Gender	2.0 (0.69-5.65)			
			Q1	2.31 (0.55)	Q1	10.1 (3.41-30.19)			
			Q2	0.50 (0.50)	Q2	1.6 (0.61-4.38)			
			Q3	0.53 (0.36)	Q3	1.7 (0.83-3.46)			
			Q4	0.06 (0.31)	Q4	1.1 (0.57-1.97)			
			Q5	-0.6 (0.63)	Q5	0.5 (0.15-1.90)			
Model 2: Age+gender+, Q1-Q10	0.94	152.6	Q6	0.53 (0.21)	Q6	1.7 (1.10-2.60)			
			Q7	0.80 (0.24)	Q7	2.2 (1.38-3.60)			
			Q8	2.36 (0.57)	Q8	10.6 (3.46-32.67)			
			Q9	0.70 (0.70)	Q9	2.0 (0.51-7.95)			
			Q10	0.42 (0.17)	Q10	1.5 (1.08-2.15)			
			Age	0.13 (0.07)	Age	1.1 (0.98-1.32)			
			Gender	0.55 (0.42)	Gender	1.7 (0.77-3.97)			
			Model 3: Age+gender+ factor 1-factor3 (case)	0.89	195.3	Factor(case)	2.76 (0.37)	Factor1(case)	15.8 (7.57-32.91)
						Factor2(case)	0.85 (0.34)	Factor2(case)	2.3 (1.19-4.58)
						Factor3(case)	1.05 (0.38)	Factor3(case)	2.9 (1.34-6.17)
Age	0.14 (0.07)	Age				1.1 (0.98-1.32)			
Model 4: Age+gender+ factor 1-factor2 (control)	0.89	170.7	Gender	0.50 (0.41)	Gender	1.7 (0.77-3.97)			
			Factor1(control)	1.26(0.16)	Factor1(control)	3.5 (2.53-4.90)			
			Factor2(control)	1.11(0.20)	Factor2(control)	3.0 (2.03-4.53)			
			Age	0.15 (0.08)	Age	1.2 (0.99-1.38)			
Model 5: Age+gender+ factor 1-factor2 (control+control)	0.91	168.4	Gender	0.71 (0.44)	Gender	2.0 (0.86-4.90)			
			Factor1(control)	2.48(0.31)	Factor1(control)	12.0 (6.48-22.4)			
			Factor2(control)	1.10(0.28)	Factor2(control)	3.0 (1.74-5.25)			

**Table 4.6-2** 各個模型 AUC (Area Under Curve) 值及與 DVSS 比較

Model (included parameters)	AUC(95CI)	Estimate of AUC difference between the tested Model and Model 1)	p-value (compared with Model 1)
Model 1: Age+Gender+DVSS	0.89(0.85-0.94)	-	-
Model 2: Age+Gender+Q1-Q10	0.94(0.91-0.98)	0.05	0.005
Model 3: Age+Gender+factor1 to factor3 (case)	0.89(0.85-0.94)	-0.0002	0.99
Model 4: Age+Gender+factor1+factor2 (control)	0.89(0.85-0.94)	0.0006	0.97
Model 5: Age+Gender+factor1+factor2 (case+control)	0.91(0.85-0.94)	0.02	0.27



## 4.7 貝氏定理應用於以DVSS篩檢社區及門診兒童

根據過去小型的研究，一般社區兒童的排尿障礙盛行率為 2%至 15%[1, 7, 9]，至於小兒泌尿門診來求診的兒童，具有排尿障礙的盛行率為 30%[10]，因此用同樣的中文文化DVSS問卷於社區及小兒泌尿門診篩檢有兒童排尿障礙的病人，同樣的分數可能會有不同具有兒童排尿障礙的機率 (probability) 及風險比 (odds ratio)，由Table 4.7-1 可見。如當於社區施行篩檢時，如果一般社區的排尿障礙盛行率為 2%，而在門診的盛行率為 30%，此時同樣的DVSS總分 8 分，在社區可能為個案的機率只有 8%，而在門診則機率增加至 63.5%。

Table 4.7-1 各個 DVSS 總分在不同盛行率地區具有兒童排尿障礙的機率 (probability) 及風險比 (odds ratio)

DVSS	Prevalence 2%		Prevalence 7%		Prevalence 15%		Clinics with 30% prevalence	
	Posterior probability	Posterior odds	Posterior probability	Posterior odds	Posterior probability	Posterior odds	Posterior probability	Posterior odds
0	0.01	0.01	0.02	0.02	0.04	0.04	0.09	0.10
1	0.01	0.01	0.02	0.02	0.04	0.04	0.10	0.11
2	0.01	0.01	0.02	0.02	0.05	0.05	0.11	0.12
3	0.01	0.01	0.03	0.03	0.06	0.07	0.14	0.16
4	0.01	0.01	0.04	0.04	0.08	0.09	0.18	0.22
5	0.02	0.02	0.05	0.06	0.12	0.13	0.24	0.32
6	0.02	0.03	0.08	0.09	0.18	0.21	0.34	0.52
7	0.04	0.04	0.14	0.16	0.27	0.38	0.48	0.91
8	0.08	0.08	0.23	0.31	0.42	0.72	0.63	1.74
9	0.15	0.17	0.39	0.63	0.60	1.48	0.78	3.58
10	0.28	0.38	0.58	1.41	0.77	3.30	0.89	8.02
11	0.48	0.93	0.77	3.42	0.89	8.01	0.95	19.46
12	0.71	2.44	0.90	9.00	0.95	21.10	0.98	51.25
13	0.87	6.97	0.96	25.72	0.98	60.30	0.99	146.45
14	0.96	21.62	0.99	79.73	0.99	186.93	1.00	453.97



## Chapter 5 討論

本研究結果發現中文化的兒童排尿障礙問卷有良好的重測信度 (ICC 值為 0.89)。但個案問卷的內部一致性 Cronbach's alpha 值較低為 0.448, 可見患有兒童排尿障礙的兒童其症狀為一個異質化的組成。因此, 當我們以排尿障礙兒童的症狀作因素分析時, 發現十個症狀可以分成三個潛在因素: 第一個為尿失禁及急尿症狀 (Q1, Q2, Q6, Q7), 第二個為排泄困難症狀 (Q3, Q4, Q8), 第三個壓力及疼痛症狀 (Q5, Q9, Q10)。比較個案組與對照組的十個題目及總分, 除了低排尿次數 (Q5) 之外, 大便, 小便及壓力事件相關的分數, 都有顯著的差異。如果以中文化兒童排尿障礙問卷量表的總分作為預測是否為有排尿障礙的個案, 由 Figure 4.4-1 及 Table 4.4-1 考量, 我們選擇 6.66 作為切點, 此時可以有 81.67% 的敏感性及 82.13% 的特異性, 而 AUC 為 0.89。最後, 當以中文化 DVSS 問卷於社區及小兒泌尿門診篩檢有兒童排尿障礙的病人, 同樣的分數會有顯著不同具有兒童排尿障礙的機率 (probability) 及風險比 (odds ratio)。

我們認為 DVSS 為一個非侵襲性的問卷量表, 當兒童的症狀分數高於切點, 而懷疑有排尿障礙時, 根據現今的診斷共識 [1], 只要臨床醫師評估及非侵襲性的尿路動力學檢查, 便可以診斷是否患有此疾病, 因此我們在設定切點時可以犧牲一些特異性來換取較好的敏感性, 在女童的 DVSS 總分, 雖然切點選擇 8 有較佳的特異性 (93.79%) 不過敏感性只有 71.05%, 因此我們選擇 >6.66 作為切點, 此時可以有 81.58% 的 sensitivity 及 77.93% 的 specificity (Table 4.4-2)。而男童以 >6.66 作為切點有較好的敏感性 (81.82%) 及特異性 (88.89%) (Table 4.4-3)。因此我們的研究結果為不管男童或女童或女生都選擇以 DVSS >6.66 作為切點。這結果和 Farhat et al 原始設計的英文問卷研究認為男童的切點設定 (9 分) 較高於女童 (6 分) 的結果不同。

由於兒童排尿障礙症狀及表現為相關多變量變數, 而且由 Cronbach's alpha 值僅有 0.448,

合理認為兒童障礙症狀的組成相當異質化，但原始作者並沒有試著將所有的症狀作分類，也無法知道這十個症狀之間的相關性，因此唯有藉著因素分析去找出隱藏在這些症狀之後的潛在變項，並依這些症狀的因素負荷將這些症狀分類，再藉由這些面向建構兒童排尿障礙的症狀表現。過去Afshar et al[5]曾經將他們的兒童排尿障礙問卷依因素分析將所有症狀分類為四：1. 尿失禁 (incontinence), 2. 急尿 (urgency), 3. 排泄阻塞症狀 (obstructive symptoms), 及 4. 大便失禁 (fecal soiling) 等症狀，不過他們並沒有就這些潛在因素作解釋及進一步解析哪些症狀屬於哪一個分類。而我們根據有排尿障礙兒童的症狀作因素分析，發現可以分成三個因素：第一個潛在因素為尿失禁及急尿症狀為膀胱逼尿肌過動的症狀，由於排尿障礙的兒童在尿路動力學檢查會有逼尿肌過動，不自主收縮造成尿急，甚至尿失禁的情形，而兒童為了避免尿失禁而會有急尿屈膝的動作。我們由因素分析發現與逼尿肌過動相關的四個症狀 (Q1 尿失禁頻率, Q2 尿失禁量, Q6 屈膝急尿及 Q7 急尿症狀) 都在第一個潛在因素的因素負荷較大，合理的我們可以把這四個症狀歸在一類。第二個潛在因素為排泄困難症狀，包含了便秘和排尿困難的症狀，便秘的小孩容易會有大便較硬，排便困難及排便次數減少等症狀，由於長期便秘的兒童會有骨盆底肌肉活性過高及放鬆不良，造成排尿時尿道括約肌無法放鬆，造成排尿時需用力，因此都在第二個潛在因素的因素負荷較大，合理的我們把 Q3 低排大便次數, Q4 排大便時需用力, 及 Q8 小便需用力等症狀有相關性可以歸在同一類症狀。最後則為，壓力及疼痛症狀包含了 Q5 排尿次數減少, Q9 解尿疼痛, 及 Q10 壓力事件，我們只能臆測小朋友遇到壓力事件時，可能會有憋尿或者排尿疼痛等症狀反應。由我們的結果和 Afshar et al[5] 的研究結果比較，兒童排尿障礙的症狀分類大致類似，只是我們把急尿和尿失禁分在同一類，而 Afshar et al[5] 的問卷並沒有包含壓力事件。而我們經由羅吉斯迴歸將個別症狀 (Q1-Q10) 帶入分析發現，Q1 (urine incontinence), Q6 (Curtsyng), Q7 (urgency), Q8 (difficulty voiding) 和 Q10 (stressful events) 預測個案的能力較好。可見雖然有些學者認為壓力事件不該被放進量表，但由我們的結果發現，兒童的

排尿障礙容易受壓力事件影響，這和過去的研究發現兒童的尿床及下尿路症狀容易受父母離異及家中變故而影響病情相同[19]。

預測性效度指的是不管個案及對照組都有已經有確切診斷，而國際小兒尿失禁醫學會認為兒童排尿障礙的診斷共識[1]，標準為臨床判斷加上尿流速，肌電圖，及殘尿量檢查，全套尿路動力學檢查的角色已經變得較不重要。因本研究的個案及對照組的診斷標準符合診斷共識，個案組為經臨床小兒泌尿科醫師評估並經過非侵襲性尿流速及膀胱超音波檢查認定有排尿障礙的病人，而個案組也選取健康，無泌尿及神經科病史，且尿流速及殘尿量正常的兒童，因此適用預測性效度。本研究結果在預測效度方面，經羅吉斯迴歸和ROC曲線分析比較五個模型 (Table 4.6-1 和 Table 4.6-2)，以模型一: age + gender + DVSS(AUC=0.89)為基準，我們發現模型二因為將個別症狀帶入模型，並依個別症狀作加權，因此其預測個案的能力較好 (AUC=0.94, AUC difference with Model 1=0.05,  $p<0.01$ )。如果以這十個問題簡化成三個潛在因素作羅吉斯迴歸作預測效度分析(模型三)，發現AUC為 0.89，其中因素一至因素三的風險比分別為 15.8 (95CI 7.57~32.91), 2.3 (95CI:1.19~4.58)，和 2.9 (95CI:1.34~6.17)，和模型一的AUC並沒有差異。而模型四以對照組兒童症狀的兩個因素作羅吉斯迴歸，因素一和因素二的風險比各為 3.5 (95CI: 2.53~4.90)及 3.0 (95CI: 2.03~4.53)，因此因素一：尿失禁及急尿症狀 (overactive bladder symptoms) 和因素二：排泄症狀 (elimination symptoms)，兩個潛在因素便可以達到很好的預測效度。但如果和模型一比較，AIC值的表現較好，但AUC為 0.89 和模型一的AUC並沒有差異。而模型五以所有個案及對照組兒童症狀的兩個因素作羅吉斯迴歸，因素一和因素二的風險比各為 12.0 (95%CI 6.48~22.37)及 3.0 (95CI: 1.74~5.24)，因此因素一：尿失禁及急尿症狀 (overactive bladder symptoms) 和因素二：排泄症狀 (elimination symptoms)，這兩個潛在因素也可以達到很好的預測效度。和模型一比較，AIC值的表現較好，AUC值為 0.91 較模型一高，但未達統計顯著差異。由以上結果決定，雖然DVSS總分是以十個症狀相同加權得到的總分，

但因為和模型二比較計算較為簡單，而且由患有兒童排尿障礙的病童症狀因素分析，也發現剛好三個潛在面向的問題各為 3 至 4 題，而因素一的風險比較高，也恰好有四題題目，分數比重較重，因此我們認為以原問卷總分作為預測個案與非個案的量表較容易為臨床所使用。

由於一般社區兒童的排尿障礙盛行率為 2%至 15%[1, 7, 9]，但是來小兒泌尿門診求診的兒童，有排尿障礙的機率為 30%[10]，因此當我們以中文化DVSS問卷原始總分於社區及小兒泌尿門診篩檢有兒童排尿障礙的病人，同樣的分數可能會有不同具有兒童排尿障礙的機率 (probability) 及風險比 (odds ratio)，此時同樣的DVSS總分八分，在社區 (2%盛行率) 可能為個案的機率只有 8%，而在門診則機率增加至 63.5%。以後的研究方向可以再研究在不同盛行率地區，需要切點的不同，並探討其敏感性及和特異性。並可以計算兒童有不同的症狀下會有排尿障礙的條件機率 (conditional probability)。

本研究有以下幾項限制：(1)本研究的問卷設計為翻譯原有已存在的問卷 (DVSS)，並沒有參考其他問卷的內容針對問卷內容再增加問卷題目，比如是否有半夜尿床，大便失禁，或者是頻尿的症狀。(2) 本研究的五位小兒泌尿科專家的內容效度及審視語意是否與原問卷是否有不同，為在問卷資料已開始著手收集之後才進行。不過此翻譯問卷已先由本人及原著作醫師先行討論及修改，而且所有的專家也都認為翻譯問卷的語意及內容並無不妥之處 (3) 我們沒有病童訪談 (interview)，針對有排尿症狀的兒童詢問是否有哪些項目是否有語意不清，無法理解或不適用於個案之處。(4)本問卷每個項目的分數為 0 到 3 分，由於大部分的分數分佈都在 0-1 分，是否需要針對 0-1 分之間，再增加項目，減少 ceiling effect 的發生需要再進一步研究。(5) 比起斷面性研究 (cross-sectional study)，由於本研究為個案對照的研究設計，而且經過年齡及性別配對，雖然具有經濟省時，並在不考慮年齡及性別的影響之下，評估此量表的預測效度的優點，但是經過羅吉斯迴歸，會消去年齡及性別等因子對於預測排尿障礙的影響。

## Chapter 6 結論

根據本研究結果，中文化兒童排尿障礙問卷量表具有良好的重測信度，並且具有良好的預測效度，由於兒童排尿障礙為一個異質化的組成，經過因素分析發現可以將所有症狀分為三類(1.尿失禁及急尿症狀, 2.排泄困難症狀, 3.疼痛及壓力症狀)。當此中文化兒童排尿障礙問卷應用於社區及門診篩檢時，應用貝氏方法可評估同樣的分數在不同盛行率區域的患病機率及風險比。



## 參考文獻

1. Chase, J., et al., *The management of dysfunctional voiding in children: a report from the Standardisation Committee of the International Children's Continence Society*. J Urol, 2010. **183**(4): p. 1296-302.
2. Farhat, W., et al., *The dysfunctional voiding scoring system: quantitative standardization of dysfunctional voiding symptoms in children*. J Urol, 2000. **164**(3 Pt 2): p. 1011-5.
3. Akbal, C., et al., *Dysfunctional voiding and incontinence scoring system: quantitative evaluation of incontinence symptoms in pediatric population*. J Urol, 2005. **173**(3): p. 969-73.
4. Tokgoz, H., et al., *Assessment of urinary symptoms in children with dysfunctional elimination syndrome*. Int Urol Nephrol, 2007. **39**(2): p. 425-36.
5. Afshar, K., et al., *Development of a symptom score for dysfunctional elimination syndrome*. J Urol, 2009. **182**(4 Suppl): p. 1939-43.
6. Koff, S.A., T.T. Wagner, and V.R. Jayanthi, *The relationship among dysfunctional elimination syndromes, primary vesicoureteral reflux and urinary tract infections in children*. J Urol, 1998. **160**(3 Pt 2): p. 1019-22.
7. Yeung, C.K., J.D.Y. Sihoe, and S.B. Bauer, *Campbell-Walsh Urology: chapter 123: Voiding Dysfunction in Children: Non-Neurogenic and Neurogenic*. 9 ed, ed. M. ALAN J. WEIN, PhD. 2007.
8. Averbeck, M.A. and H. Madersbacher, *Constipation and LUTS - how do they affect each other?* Int Braz J Urol, 2011. **37**(1): p. 16-28.
9. Hoebeke, P., et al., *Assessment of lower urinary tract dysfunction in children with non-neuropathic bladder sphincter dysfunction*. Eur Urol, 1999. **35**(1): p. 57-69.
10. Hoebeke, P., et al., *One thousand video-urodynamic studies in children with non-neurogenic bladder sphincter dysfunction*. BJU Int, 2001. **87**(6): p. 575-80.
11. Streiner, D.L. and G.R. Norman, *Health Measurement scales A practical guide to their development and use*. 2008.
12. Skinner, H.A. and H. Lei, *Differential weights in life change research: useful or irrelevant?* Psychosom Med, 1980. **42**(3): p. 367-70.
13. Cronbach, L.J., *Coefficient alpha and the internal structure of tests*. Psychometrika, 1951. **16**(3): p. 38.
14. Nunnally, J.C., *Psychometric theory*. 2nd ed. 1978, New York: McGraw-Hill.
15. Ponterotto, J.G. and D.E. Ruckdeschel, *An overview of coefficient alpha and a reliability matrix for estimating adequacy of internal consistency coefficients with psychological research measures*. Percept Mot Skills, 2007. **105**(3 Pt 1): p. 997-1014.
16. S.Everitt, B. and G. Dunn, *Applied Multivariate Data Analysis*. 2 ed. 2001.
17. Neveus, T., et al., *The standardization of terminology of lower urinary tract function in children and adolescents: report from the Standardisation Committee of the International Children's Continence Society*. J Urol, 2006. **176**(1): p. 314-24.
18. DeLong E. R., D.D.M., Clarke-Pearson D.L, *Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach*. Biometrics, 1988. **44**(3): p. 9.
19. Devlin, J.B., *Prevalence and risk factors for childhood nocturnal enuresis*. Ir Med J, 1991. **84**(4): p. 118-20.