

修訂版學步期自閉症檢核表 在不同場域篩檢自閉症類群 障礙症幼兒的正確率

洪楹佳

高雄醫學大學心理學系
碩士生

陳奕臻

彰化基督教兒童醫院
兒童發展中心
臨床心理師

吳進欽

高雄醫學大學心理學系
中和紀念醫院臨床醫學研究部
教授 / 兼任研究員

修訂版學步期自閉症檢核表 (Modified Checklist for Autism in Toddlers, M-CHAT) 為國際上常見的自閉症類群障礙症 (autism spectrum disorder, ASD) 幼兒篩檢工具。然而, M-CHAT 在臺灣的研究不多。因此, 本研究探究 M-CHAT 在臺灣的不同場域 (如: 基層診所與綜合醫院) 篩檢 ASD 幼兒的正確率。18-29 個月的幼兒共 272 名, 包括: ASD 幼兒 74 名、發展遲緩 (developmental delay, DD) 幼兒 87 名與一般發展 (typically developing, TD) 幼兒 111 名, 由主要照顧者填寫 23 題「是 / 非」選項的 M-CHAT。以卡方檢定 (chi-squared test) 比較 ASD、DD 與 TD 幼兒在 M-CHAT 題目失敗率的差異, 單因子變異數分析 (analysis of variance, ANOVA) 檢驗三組幼兒在總分及不同簡版分數的差異。然後, 以訊號偵測理論 (signal detection theory, SDT) 決定 M-CHAT 的切截分數, 並以判別函數分析 (discriminant function analysis, DFA) 找出 M-CHAT 的結構矩陣, 決定關鍵題目。在綜合醫院場域使用時, M-CHAT 篩檢 ASD 與 DD 幼兒, 以任 4 題失敗作為切截分數, 敏感度與特異度皆為 .72。使用不同簡版時的敏感度為 .62- .72, 特異度為 .85- .91, 簡版 14 題 (Brief 14) 篩檢率正確率最佳。在基層診所場域使用時, M-CHAT 篩檢 ASD 與 TD 幼兒, 以任 3 題失敗作為切截分數, 敏感度為 .81, 特異度為 .83。使用不同簡版時的敏感度為

* 本文以吳進欽為通訊作者 (jinnchin@mail2000.com.tw)。

** 致謝: 本文的完成, 要感謝所有參與研究的兒童與家長, 以及科技部經費支持 (計畫編號: MOST103-2628-H-037-001-MY2)。

.62- .72，特異度為 .95- .99，一樣是簡版 14 題篩檢正確率最佳。M-CHAT 用在綜合醫院場域篩檢 ASD 幼兒正確率尚可，用在基層診所場域篩檢 ASD 幼兒正確率良好。本研究結果支持 M-CHAT 可以在臺灣用於 ASD 幼兒的早期篩檢，特別是在基層診所場域。然而，在不同場域使用 M-CHAT，需要使用不同的切截分數。

關鍵詞：自閉症類群障礙症、修訂版學步兒自閉症檢核表、特異度、敏感度

緒論

自閉症類群障礙症 (Autism Spectrum Disorder, ASD) 是先天的神經發展障礙，患者出現兩大類症狀，分別是：(A) 社交溝通與社交互動的缺損；(B) 侷限、重複的行為、興趣或活動模式 (American Psychiatric Association [APA], 2013)。當代研究指出 ASD 盛行率約 1-2% (APA, 2013; Christensen et al., 2016)，學齡前 ASD 兒童盛行率，臺灣約 0.2% (Lai, Tseng, Hou, & Guo, 2012)，西方國家約 0.6% (Chakrabarti & Fombonne, 2005)，顯示臺灣可能有不少學齡前 ASD 兒童沒有被診斷出來接受早期療育的服務。過去學界已經有一些研究針對 2 歲前 ASD 幼兒進行篩檢 (Sturmer, Howard, Bergmann, Stewart, & Afarian, 2017)，或診斷 (Barbaro & Dissanayake, 2017; Guthrie, Swineford, Nottke, & Wetherby, 2013)。然而，ASD 早期篩檢的研究，臺灣比較少見，特別是針對幼兒的篩檢。ASD 的早期診斷有助早期介入，幫助家長因應壓力與處理兒童的問題。因此，ASD 篩檢工具正確率的探究不容小覷。ASD 篩檢工具分為兩個層次，層次一是在一般發展 (typically developing, TD) 兒童中篩檢出疑似發展異常的患者，通常是家長填寫問卷，一般用在基層診所 (primary clinics) 場域，如：社區的小診所；層次二篩檢針對高風險族群，如：發展遲緩 (developmental delay, DD) 兒童，鑑別 ASD 與 DD 兒童，一般在綜合醫院 (general hospitals) 場域使用，如：地區醫院或區域醫院。層次二篩檢工具需要較長時間進行評估，通常由專業人員使用或執行。

目前有幾項工具被公認為 ASD 標準化評估工具，在診斷時扮演重要角色，分別為：「自閉症診斷會談問卷修訂版 (Autism

Diagnostic Interview-Revised, ADI-R; Lord, Rutter, & Le Couteur, 1994)」、「自閉症診斷觀察量表 (Autism Diagnostic Observation Schedule, ADOS; Lord, Rutter, DiLavore, & Risi, 1999)」，及「自閉症診斷觀察量表第二版 (Autism Diagnostic Observation Schedule Second Edition, ADOS-2; Lord et al., 2012)」。但是，這三項工具，施測過程繁瑣、耗費時間長，執行者需要參加費用不低的工作坊完成訓練，不論在基層診所或綜合醫院場域要使用這三項工具都有困難。因此，學界發展出不少篩檢工具，用來快速篩檢 ASD 兒童，協助進行診斷。學界針對 ASD 篩檢工具進行回顧 (吳進欽、姜忠信、侯育銘, 2013; 姜忠信、宋維村, 2005)，家長問卷部分，如：一歲期間卷 (First Year Inventory, FYI)、學步期自閉症檢核表 (Checklist for Autism in Toddlers, CHAT)、修訂學步期自閉症檢核表 (Modified Checklist for Autism in Toddlers, M-CHAT)；行為觀察部分，如：兩歲期自閉症篩檢工具 (Screening Tool for Autism in Two-Year-Olds, STAT)、溝通與象徵行為量表：發展表現 (Communication and Symbolic Behavior Scales: Developmental Profile, CSBS DP) 行為樣本 (behavior sample)。FYI 有 63 題，題目數過多，不符合篩檢的用途；M-CHAT 及 CHAT 題數較少，分別為 23 題及 14 題，比較適合用來篩檢。STAT 是設計用來篩檢 24-36 個月 ASD 幼兒的互動式篩檢工具，共有 12 題，需要由受過訓練之臨床工作者執行此評估作業。姜忠信、吳進欽、劉俊宏與侯育銘 (2012) 修改 STAT 部分題目，稱為「臺灣版兩歲期自閉症篩檢工具 (Taiwan Version of Screening Tools for Autism in Two-Year-Olds, T-STAT)」，一樣用來篩檢 24-36 個月 ASD 幼兒；CSBS DP 行為樣本用在篩檢

13-27 個月的 ASD 幼兒，共 29 題，但是計分方式較為複雜。相較於行為觀察的篩檢工具，家長問卷具方便性且省時又經濟，在臺灣臨床工作者負擔相當重的狀況下，具信、效度及篩檢正確率的家長問卷更是重要。再者，家長問卷的另一個優點是不會受限於場域而有無法執行的問題。中文版的篩檢工具不多，M-CHAT 中文版經過官方認證，在臺灣亦被許多臨床工作者使用，是值得推廣的工具。然而，西方發展的 ASD 篩檢工具，在臺灣使用需要留意文化差異 (Stewart & Lee, 2017)。目前臺灣使用 M-CHAT 都是參考國外的研究結果，有必要在臺灣驗證 M-CHAT 的切截分數、敏感度及特異度。

M-CHAT (Robins, Fein, Barton, & Green, 2001)，是國際上研究結果相當豐碩的 ASD 幼兒篩檢工具，已經在許多不同文化被驗證，如：沙烏地阿拉伯 (Seif Eldin et al., 2008)、斯里蘭卡 (Perera, Wijewardena, & Aluthwelage, 2009)、西班牙 (Canal-Bedia et al., 2011)、南韓 (Seung et al., 2015)、伊朗 (Samadi & McConkey, 2015)、土耳其 (Kondolot et al., 2016)、埃及 (Mohamed et al., 2016)、泰國 (Srisinghasongkram, Pruksananonda, & Chonchaiya, 2016) 及日本 (Inada, Koyama, Inokuchi, Kuroda, & Kamio, 2011; Kamio et al., 2014) 等。但是，臺灣探究 M-CHAT 篩檢 ASD 幼兒正確率的研究不多。回顧臺灣 ASD 篩檢的研究，可以發現 ASD 幼兒的篩檢工具相當不足，這將會妨礙早期篩檢工作。M-CHAT 官網已經有可以使用的中文版，值得在臺灣進行篩檢正確率的探究。雖然 Robins 等人 (2014) 針對 M-CHAT 繼續修訂，刪減成 20 題，稱為「修訂學步期自閉症檢核表：追蹤問題修訂版 (Modified Checklist for Autism in Toddlers, Revised with

Follow-up, M-CHAT-R/F)」。不過，目前學界研究結果，還是以 M-CHAT 為主。

Robins 等人 (2001) 提出 ASD 篩檢工具在早期篩檢的一些限制，如：大多是針對學齡兒童、評估指標 3 歲以前的 ASD 幼兒很少出現等。因此，Robins 等人修改 CHAT (Baron-Cohen, Allen, & Gillberg, 1992) 的題目，刪除其中的 5 題行為觀察，只保留 9 題家長訪談，並增加題目為 30 題。最後，30 題中再刪除 8 題缺乏鑑別力或家長容易誤解的題目，另外增加 1 題，成為 23 題的 M-CHAT，並以 23 題失敗 3 題作為切截分數 (cutoff)。除 23 題失敗 3 題的指標外，Robins 等人使用判別函數分析 (Discriminant Function Analysis, DFA)，決定 6 題關鍵題失敗 2 題當作切截分數，此 6 題分別為：2、對其它小孩感興趣；7、分享手指指示；9、展示；13、模仿；14、叫名反應，及 15、跟隨手指指示。參與研究的 1293 名 18-30 個月大幼兒，其中 1122 名來自基層診所場域，171 名來自早期介入機構，原本 30 題中失敗 3-5 題的幼兒，進行電話訪談追蹤。結果發現 23 題失敗 3 題當作切截分數，敏感度 (sensitivity) .97，特異度 (specificity) .95，正向預測值 (positive predictive value, PPV) .36，負向預測值 (negative predictive value, NPV) .99，經電話追蹤，PPV 明顯增加，但 NPV 不變。使用 6 題關鍵題失敗 2 題當作切截分數，敏感度與特異度和全部 23 題差別不大，但 PPV 較高，經電話追蹤，PPV 提高至 .79，表一為 M-CHAT 相關研究結果。然而，研究指出 ASD 與 DD，及 ASD 與 TD 的鑑別指標不一樣 (Osterling, Dawson, & Munson, 2002; Veness et al., 2012)，Robins 等人的研究，納入早療機構的幼兒，可能因家長對小孩狀況的態度與想法而出現偏差，這兩個篩檢指標的合宜性，

及篩檢正確率，需要透過招募不同對照組持續驗證。近幾年，M-CHAT 出現最佳 7 題 (Best 7) 替代關鍵 6 題的計分方式 (Robins et al., 2010)。最佳 7 題為：2、對其他小孩感興趣；5、假裝遊戲；7、分享手指指示；9、

展示；14、叫名反應；15、跟隨手指指示及 20、聽力問題。不過，最佳 7 題篩檢正確率的研究不多。

M-CHAT 在基層診所場域進行篩檢，以 TD 組為對照組的研究。Robins (2008)

表一 修訂版學步期自閉症檢核表敏感度、特異度的研究

作者及年份	研究對象	篩檢標準	敏感度	特異度
Robins et al. (2001)	18-30 個月，1293 名 (1122 名低風險，171 名高風險)	23 題失敗 3 題 關鍵 6 題失敗 2 題	.97 .95	.95 .98
Chlebowski et al. (2013)	16-30 個月，18989 名 TD	23 題失敗 3 題	不詳 不詳	不詳 不詳
Srisinghasongkram et al. (2016)	18-48 個月，841 名幼兒 (732 名低風險，109 名高風險)	23 題失敗 3 題或 關鍵 6 題失敗 2 題	1	.29
Topçu et al. (2018)	16-38 個月，486 名定期健康檢查幼兒	23 題失敗 3 題或 關鍵 6 題失敗 2 題	.60	.96
Eaves et al. (2006)	17-48 個月，84 名，54 名 ASD	關鍵 6 題失敗 2 題 19 題失敗 3 題	.77 .92	.43 .27
Snow and Lecavalier. (2008)	18-48 個月，56 名，43 名 ASD	23 題失敗 3 題 關鍵 6 題失敗 2 題	.88 .70	.38 .38
Salisbury et al. (2018)	16-30 個月，271 名，疑似發展問題的幼兒，其中 173 名 ASD	23 題失敗 3 題或 關鍵 6 題失敗 2 題	.78	.54
Wong et al. (2004)	13-86 個月，276 名 (實際納入分析的只有 212 位)，其中 87 名 ASD	23 題失敗 3 題 23 題失敗 6 題 關鍵 7 題失敗 2 題	.95 .84 .93	.54 .75 .77
Koh et al. (2014)	18-30 個月，173 名高風險幼兒	23 題失敗 3 題 關鍵 6 題失敗 2 題 最佳 7 題失敗 2 題	.89 .75 .81	.59 .78 .78
Toh et al. (2018)	15-36 個月，50 名 ASD 15-20 個月；21-26 個月； 27-36 個月	23 題失敗 3 題或 關鍵 6 題失敗 2 題	.36 .18; .55; .64	1 皆為 1
Lung et al. (2011)	36 個月以下，超過 1600 名 ASD	23 題失敗 3 題 關鍵 6 題失敗 2 題	不詳 不詳	不詳 不詳
Wong et al. (2018)	18-47 個月，236 名，其中 113 名 ASD，123 名 DD	23 題失敗 3 題 23 題失敗 4 題 關鍵 6 題失敗 2 題 最佳 7 題失敗 2 題 簡版 14 題失敗 3 題	.88 .77 .46 .53 .71	.53 .72 .94 .87 .82

的研究參與者為來自定期健康檢查 (well-child visit) 的 4797 名 14-27 個月大幼兒，但僅 41 名接受診斷評估，其中 21 名為 ASD。結果發現 ASD 在 M-CHAT 的失敗題數顯著較多，不過 Robins 的研究沒有報告敏感度。Chlebowski、Robins、Barton 與 Fein (2013) 的研究參與者為 18989 名 16-30 個月的 TD 幼兒，篩檢結果約 9.1% 幼兒超過切截分數，需要進行電話追蹤。在 92 名 ASD 中，23 題失敗 3 題可以篩檢出 91 名 (約 99%)，但是關鍵題目只能篩檢出 78 名 (約 85%)。之後，Srisinghasongkram、Pruksananonda 與 Chonchaiya (2016) 在泰國招募 841 名 18-48 個月大的幼兒，篩檢結果約 60% 幼兒需進行電話追蹤，全部題數的敏感度 (.91) 優於關鍵題目 (.76)。Chlebowski 等人與 Srisinghasongkram 等人的研究，皆發現使用所有題目進行篩檢是比較好的選擇。後續 M-CHAT 在基層診所場域的研究結果，一樣發現特異度相當理想，但是敏感度不是很穩定 (Topçu, Ulukol, Öner, imek, OrhonBakan, 2018)。

M-CHAT 在綜合醫院場域篩檢 ASD 兒童，Eaves、Wingert 與 Ho (2006) 針對 84 名 17-48 個月的兒童進行篩檢，其中 54 名為 ASD。結果發現以關鍵 6 題失敗 2 題作為切截分數，敏感度 .77，特異度 .43；使用 19 題失敗 3 題作為切截分數，敏感度提升但特異度下降。後續 Snow 與 Lecavalier (2008) 針對 56 名 18-48 個月大的兒童進行探究，其中 43 名為 ASD。結果發現 23 題失敗 3 題作為切截分數，敏感度 .88，特異度 .38。此外，偵測者運作特徵 (receiver operating characteristic, ROC) 曲線下面積 .70；關鍵 6 題失敗 2 題作為切截分數，敏感度較低，特異度相同，ROC 曲線下面積 .65。一篇

大樣本病歷回顧研究，Salisbury、Nyce、Hannum、Sheldrick 與 Perrin (2018) 的研究參與者為 271 名 16-30 個月疑似發展問題的幼兒，其中 173 名為 ASD，M-CHAT 篩檢結果敏感度 .78，特異度 .54。三篇研究結果顯示 M-CHAT 用在綜合醫院場域，敏感度不錯，但是特異度不佳，這有幾個可能的原因，1、原有的篩檢指標比較適合用在基層診所場域進行篩檢，而非在綜合醫院場域進行篩檢。2、綜合醫院場域的研究參與者其生理年齡大於 M-CHAT 原本設定的年齡範圍，導致篩檢結果不理想。

篩檢工具對不典型 (atypical) 行為的定義，可能存在文化差異，影響工具在不同文化的適用性 (Wallis & PintoMartin, 2008; Stewart & Lee, 2017)。M-CHAT 雖然在許多文化中被使用，但是華人文化的研究比較少。Wong 等人 (2004) 在香港把 M-CHAT 翻譯成中文版本，研究對象為 276 名 13-86 個月大的兒童 (實際納入分析的只有 212 位)，其中 87 名為 ASD。結果發現 23 題失敗 3 題，敏感度 .95、特異度 .54，PPV .59；23 題失敗 6 題，敏感度下降，但特異度與 PPV 提高，ROC 曲線下面積 .91。Wong 等人發現 7 題關鍵題，分別是：2、社交興趣；5、假裝遊戲；7、分享式手指指示；9、展示；13、模仿；15、跟隨手指指示，及 23、社會參照。7 題關鍵題失敗 2 題，敏感度 .93，特異度 .77，PPV .74，ROC 曲線下面積 .92。Koh 等人 (2014) 在新加坡的綜合醫院場域篩檢 173 名 18-30 個月大的高風險幼兒，發現 23 題失敗 3 題，敏感度 .89、特異度 .59、PPV .49、NPV .92，ROC 曲線下面積 .84；若使用關鍵 6 題與最佳 7 題失敗 2 題，敏感度與 NPV 下降、特異度與 PPV 提高，ROC 曲線下面積相當。最近，馬來

西亞基層診所場域的研究，50 名 15-36 個月大的 ASD 幼兒，僅 18 位 M-CHAT 分數超過切截分數（任 3 題失敗或關鍵 6 題失敗 2 題），敏感度 .36，特異度 1（Toh, Tan, Lau, & Kiyu, 2018）。然而，Toh 等人的研究，以幼兒在 M-CHAT 兩個指標任一為陽性來計算敏感度，無法了解哪一個指標篩檢正確率較為理想，且此方式的敏感度相對較高；再者，篩檢結果為陽性的 TD 幼兒沒有進行評估也是一個缺點。Wong 等人的研究則存在將 DD 與 TD 合併成對照組及篩檢年齡範圍太大的問題，Koh 等人的研究則沒有明確交代臨床樣本的屬性。因此，這三篇研究結果的解釋需要謹慎。

M-CHAT 的研究結果相當豐碩，關於其篩檢正確率的研究，Yuen、Penner、Carter 與 Ungar（2018）的後設分析（meta-analysis），回顧 13 篇文獻，包含層次一與層次二的篩檢研究。結果發現 M-CHAT 的敏感度平均 .83，95% 信賴區間 .75- .90，特異度平均 .51，95% 信賴區間 .41- .61。最近，一樣是透過後設分析，Sánchez-García、Galindo-Villardón、Nieto-Librero、Martín-Rodero 與 Robins（2019）回顧 18 篇 ASD 層次一篩檢工具的文章，主要的篩檢工具為 M-CHAT。結果發現敏感度平均 .72，95% 信賴區間 .61- .81，特異度介於平均 .98，95% 信賴區間 .97- .99。兩篇後設分析的研究結果，發現 M-CHAT 的敏感度還算穩定，但是特異度因為場域的差異而出現極大的差異，這個結果更彰顯針對不同場域探究篩檢正確率的必要性與重要性。

M-CHAT 在臺灣的研究，Lung、Chiang、Lin 與 Shu 等人（2011）針對臺灣超過 1600 名 3 歲以下幼兒進行兩階段篩檢。第一階段使用臺灣出生世代研究（Taiwan

Birth Cohort Study, TBCS）的發展評估工具，第二階段使用 M-CHAT 作為 ASD 的分類。研究結果發現兩個工具分類結果存在相關，有 292 名（18%）幼兒 23 題失敗 3 題，24 名（1.5%）6 題關鍵題失敗 2 題。不過，Lung 等人並沒有針對參與研究的幼兒進行正式診斷，無法知道 M-CHAT 的敏感度與特異度為何。之後，Wong 等人（2018）進行臨床樣本研究，236 名 18-47 個月兒童，其中 113 名為 ASD，123 名為 DD。23 題失敗 3 題，敏感度 .88、特異度 .53、PPV .63、NPV .82，ROC 曲線下面積 .82。Wong 等人還發現簡版 14 題（brief 14）失敗 3 題相較於其他不同簡版，有較好的篩檢正確率，敏感度 .71、特異度 .82、PPV .78、NPV .65，ROC 曲線下的面積 .83。Wong 等人的研究顯示以 23 題失敗 4 題作為指標，敏感度與特異度較為理想。不過，該研究沒有分開探究 18-30 個月與 31-47 個月，不確定 M-CHAT 用在 30 個月以下幼兒的篩檢正確率。

除了 M-CHAT 外，臺灣 ASD 幼兒早期篩檢的研究，尚有以下工具：（1）克氏行為量表（Clancy Behavior Scale, CBS; 吳進欽、朱慶琳、侯育銘、姚淑芬，2014；謝清芬、宋維村、徐澄清，1983）；（2）學步期自閉症篩檢問卷（Toddler Autism Screening Questionnaire, TASQ; Tsai, Soong, & Shyu, 2012）；（3）中文版泛自閉症篩檢量表（Autism Spectrum Rating Scales Chinese Version, ASRS-C; 阮麗敏、林玲伊，2015），三個工具當中，CBS 是比較廣泛被使用且研究結果較多的工具。然而，CBS 為三點量表，家長會有 1 分的反應心向（吳進欽等人，2014），若有強迫選擇的家長問卷，可以補足 CBS 的缺失。過往關於 M-CHAT 的研究，有下列的發現：第一、18-30 個月的篩檢正

確率高於 31-48 個月的篩檢正確率 (Koh et al., 2014; Salisbury et al., 2018)，且不同年齡層篩檢指標不一樣 (Barbaro & Dissanayake, 2010)。因此，需要針對 Wong 等人 (2018) 的研究進行驗證。第二、過去 M-CHAT 的研究，對照組將 TD 與 DD 合併 (e.g., Wong et al., 2004)；然而，ASD 早期症狀的研究，指出 ASD 與 DD，ASD 與 TD 的區辨指標不一樣 (Barbaro & Dissanayake, 2013; Osterling et al., 2002)。因此，需要分別以 TD 或 DD 當作對照組，探究切截分數與篩檢正確率。因此，本研究將針對上述問題進行修正，縮小研究參與者年齡範圍，所有參與研究的幼兒，生理年齡介於 18-29 個月，對照組則分別是 DD 與 TD 幼兒，並依據精神醫學診斷及統計手冊第五版 (diagnostic and statistical manual of mental disorders, 5th edition, DSM-5; APA, 2013) 進行臨床樣本的診斷，探究 M-CHAT 篩檢 ASD 幼兒的正確率。

本研究的主要目的如下：第一、比較 ASD、DD 與 TD 幼兒在 M-CHAT 分數的差異。第二、比較 ASD、DD 與 TD 幼兒在 M-CHAT 全部 23 題、關鍵 6 題、最佳 7 題及簡版 14 題等分數的差異。第三、探究 M-CHAT 鑑別 ASD 與 DD 幼兒、ASD 與 TD 幼兒的切截分數、敏感度、特異度、PPV 與 NPV。

研究方法

(一) 研究參與者

本研究資料來自於通過區域教學醫院人體試驗委員會審核的研究計畫 (編號為：CYCH-IRB101022; CYCH-IRB102045)。275 名生理年齡 18-29 個月大的 ASD、DD 與 TD 幼兒參與本研究。所有研究參與者的

家長需能自行閱讀中文並填寫問卷，且研究參與者皆無明顯的感官、嚴重動作發展障礙。排除三位外籍配偶因對於問卷理解不佳後，最後招募的研究參與者為 272 名 18-29 個月大的幼兒，其中 ASD 幼兒 74 名、DD 幼兒 87 名，TD 幼兒 111 名。研究參與者於 2016 年 1 月至 2019 年 5 月經由下列方式招募：

1、自閉症類群障礙症組 (ASD 組)

家長因擔心小孩發展遲緩問題帶幼兒至醫院就診，由醫師或其他專業人員告知研究訊息；或早療社工人員在社區或基層診所場域進行發展遲緩篩檢，篩檢未通過者，由社工人員告知家長研究訊息。ASD 幼兒診斷由資深兒童精神科醫師或博士學位的兒童臨床心理師，依據家長主訴、發展史、行為觀察，及參考 ADOS 的評估結果，根據 DSM-5 的準則給予診斷。所有 ASD 幼兒符合症狀 A 三項及 B 兩項。

2、發展遲緩組 (DD 組)

家長因擔心小孩發展遲緩問題帶幼兒至醫院就診，由醫師或其他專業人員告知研究訊息；或早療社工人員在社區與基層診所場域進行發展遲緩篩檢，篩檢未通過者，由社工人員告知家長研究訊息。DD 幼兒的診斷依據穆林早期學習量表 (Mullen Scale of Early Learning, MSEL; Mullen, 1995) 的評估結果，在視覺接收、精細動作、語言理解及語言表達等四領域，有任一領域的 T 分數低於 40，或整體發展分數低於 85。此外，DD 幼兒的家長沒有擔心小孩有互動方面的問題。

3、一般發展組 (TD 組)

透過研究者在網路張貼宣傳研究計畫的訊息，由家長主動聯繫研究者。願意參加研究的家長，由相關人員說明研究內容及資訊，進一步安排評估時間。TD 幼兒在 MSEL

評估結果沒有出現落後，且家長亦沒有擔心小孩的互動與情緒問題。

(二) 研究工具

1、穆林早期學習量表 (Mullen Scales of Early Learning, MSEL; Mullen, 1995)

MSEL 為國際相當常用的標準化發展測驗，評估出生至 68 個月的兒童，包含五項分量表：粗動作、精細動作、視覺接收、語言理解及語言表達；後面四項分量表為認知發展量表，可以組合成發展能力，稱早期學習組合標準分數 (early learning composite standard scores)。分量表及早期學習組合標準分數的內部一致性係數之中位數均超過 .75，信度相當理想。MSEL 與貝萊嬰兒發展量表 (Bayley Scales of Infant Development; Bayley, 1969) 之心智發展與動作發展能力，共享變異數為 28-35%，同時效度良好。MSEL 在本研究中主要用途為評估幼兒的發展能力。

2、自閉症診斷觀察量表 (Autism Diagnostic Observation Schedule, ADOS; Lord et al., 1999)

ADOS 是由 Lord 等人 (1999) 發展的半結構化 ASD 行為觀察量表，透過遊戲互動方式評估兒童的溝通、社會互動、遊戲能力及同一性等行為，為國際公認的 ASD 標準化測驗。ADOS 共有四個模組 (modules)，針對不同語言表達能力選擇適當模組施測；本研究 ASD 組與 DD 組幼兒皆使用模組一。ADOS 依據溝通障礙、社會互動障礙，及溝通合併社會互動障礙分數作為分類依據，得分越高代表越符合自閉症特徵。ADOS 模組一的溝通障礙、社會互動障礙，及溝通合併社會互動障礙，自閉症切截分數分別為 4 分、7 分與 12 分；非典型自閉症切截分數分別為

2 分、4 分與 7 分，ADOS 模組一評分者信度平均達 91.5%，且所有題目的評分者信度皆超過 80%；在鑑別 ASD (包含自閉症及非典型自閉症) 與其他非自閉症的敏感度 .97，特異度 .94，信度與正確率相當理想。ADOS 需由參加工作坊取得認證資格的人員進行施測。本研究的 ADOS 由已經取得認證的通訊作者執行。

3、修訂版學步期自閉症檢核表 (Modified Checklist for Autism in Toddlers, M-CHAT; Robins et al., 2001)

M-CHAT 為一份家長填寫的問卷，共有 23 題是非題，題目分別是：1、肢體接觸；2、社交興趣；3、粗大動作一爬；4、捉迷藏；5、假裝遊戲；6、要求式手指指示；7、分享式手指指示；8、正確玩玩具；9、展示；10、眼神注視；11、聽覺敏感；12、社會性微笑；13、模仿；14、叫名反應；15、跟隨手指指示；16、粗大動作一走；17、跟隨眼神注視；18、怪異手勢動作；19、吸引注意；20、聽力問題；21、語言理解；22、眼神接觸不佳；23、社會參照。「是」表 0 分；「否」表 1 分，其中有 4 題反向題 (11、18、20、22)，分數越高代表越符合自閉症特徵。

(三) 研究程序

有意願參與研究的家長，研究前由研究人員與家長約定好時間，並在約定時間前來進行評估，評估地點為大學心理系的兒童遊戲室或教學醫院的心理測驗室。正式進行評估之前，研究人員會再說明一次研究內容，並請家長填寫知情同意書；等待家長填寫完知情同意書才開始進行評估。每位幼兒皆需要進行 MSEL 評估，僅有 ASD 組與 DD 組幼兒需要進行 ADOS 的評估。幼兒來到研究地點後，會先給予時間適應環境，然後開

始進行評估，通常會先進行 MSEL 然後進行 ADOS。不過，ADOS 與 MSEL 的施測順序可以視幼兒情況作調整。若是幼兒較為害羞的情況，因為 ADOS 的材料都是玩具且較有吸引力、不需要坐在椅子上就可以進行，就會先進行 ADOS 評估，然後進行 MSEL。幼兒進行評估時，家長在一旁陪伴幼兒並同時填寫問卷，主要是安撫幼兒；家長填寫的問卷包含親職壓力、身心健康問卷、幼兒適應行為及問題行為等，M-CHAT 為其中一份問卷。除此之外，為確保 ADOS 與 M-CHAT 結果之間的獨立性，ADOS 施測者只知道幼兒為研究參與者，對於家長主訴的問題與 M-CHAT 結果並不知情，確保 M-CHAT 與 ADOS 兩者的獨立性。因為研究人力不足，部份幼兒的臨床診斷與 ADOS 施測者會是同一位，但有任何疑慮會以團隊形式確認兒童的診斷。由於本研究中 ADOS 結果與臨床診斷，並不需要進行兩者一致性的檢驗，對研究結果影響不大。若研究人力足夠的情況下，未來可以規劃臨床診斷與 ADOS 施測者獨立評估。

（四）統計分析

使用 SPSS 套裝軟體進行分析。單因子變異數分析 (analysis of variance, ANOVA) 檢驗三組研究參與者各項基本資料，以卡方檢定 (chi-squared test) 檢驗三組幼兒男女比例，獨立樣本 t 檢定 (independent-sample t-test) 檢驗 ASD 組與 DD 組自閉症症狀。首先，以內部一致性 (Cronbach's α) 分析 M-CHAT 的信度，再以 ADOS 作為效標，以皮爾森積差相關分析 (Pearson product-moment correlation coefficient) 探究臨床參與者 M-CHAT 總分與 ADOS 的相關。針對研究目的，以卡方檢定分析三組研究參與

者在 M-CHAT 各題表現差異，考慮 α 值膨脹，取 $p < .002 (.05/23)$ ；單因子變異數分析檢驗三組研究參與者在 M-CHAT 全部 23 題、關鍵 6 題、最佳 7 題及簡版 14 題等分數的差異，並以雪費法 (Scheffe) 進行事後比較。判別函數分析 (discriminant function analysis, DFA) 分別以 ASD 組與 DD 組，及 ASD 組與 TD 組進行檢驗，M-CHAT 題目為預測變項，用於預測是否區辨 ASD 與 DD 組或 ASD 與 TD 組。利用判別函數分析取得 M-CHAT 所有題目的結構矩陣，找出具區辨能力的關鍵題目。若該題目判別函數係數值越大，代表該題目越能區辨 ASD 與 DD 組或 ASD 與 TD 組。最後，以訊號偵測理論 (signal detection theory, SDT)，即偵測者運作特徵 (receiver operating characteristic, ROC) 曲線計算 ASD 組與 DD 組及 ASD 組與 TD 組的切截分數，檢驗敏感度、特異度。

研究結果

三組研究參與者皆以 MSEL 評估其發展狀況，並參考 Rogers、Hepburn、Stackhouse 與 Wehner (2003) 的作法，將視覺接收、精細動作、語言理解及語言表達四個領域的原始分數所對應之年齡當量 (age equivalent) 相加的平均，作為心理年齡 (mental age, MA)。社經地位則以父母接受正式教育之年數當作指標：國小畢業 6、國中畢業 9、高中/職畢業 12、專科畢業 14、大學畢業 16、研究所畢業 18；其中一名幼兒因為父母離異，故缺少一名母親教育年數的資料。

以單因子變異數分析檢驗三組研究參與者各項基本資料 (見表二)。三組幼兒的生理年齡、心理年齡、母親教育年數及父親教育年數皆達顯著差異。以雪費法事後比較發

現，生理年齡部分，ASD 組與 TD 組顯著高於 DD 組，但 ASD 組與 TD 組沒有顯著差異；社經地位，母親教育、父親教育部分，TD 組顯著高於 ASD 組與 DD 組，ASD 組與 DD 組沒有顯著差異；心理年齡，TD > DD > ASD。以卡方檢定檢驗三組幼兒男女比例 ($\chi^2 = 33.36, p < .001$)，結果達顯著差異，ASD 組女生人數顯著較少。以獨立樣本 t 檢定檢驗 ASD 組與 DD 組自閉症症狀達顯著差異，ASD 幼兒症狀較為嚴重，支持臨床樣本診斷的可靠性。

內部一致性檢驗 M-CHAT 的信度，內部一致性係數為 .82，逐題刪除後的內部一致性係數介於 .80- .84，顯示 M-CHAT 具良好內部一致性（見表三）。然而，第 1 題、第 3 題、第 11 題、第 16 題及第 18 題與 M-CHAT 總分

的相關係數低於 .20。ADOS 分數作為效標，皮爾森積差相關分析 ASD 組、DD 組與臨床組（ASD 合併 DD）在 ADOS 三項指標的相關，ASD 組相關係數介於 .22- .44，M-CHAT 總分與 ADOS 總分、ADOS 互動達顯著相關 ($p < .001$)；DD 組相關係數介於 .03- .16，皆未達顯著相關；臨床組（ASD 合併 DD）相關係數介於 .54- .61，皆達顯著相關。依據 Dancey 與 Reidy (2007) 的建議，相關係數大於或等於 .70，屬高度相關，.40- .70 屬中度相關，.40 以下屬低度相關，本研究結果顯示 M-CHAT 具高度內部一致性，M-CHAT 與臨床組（ASD 合併 DD）的 ADOS 具中度效標關聯效度。

以卡方考驗檢定 ASD 組、DD 組與 TD 組在 M-CHAT 表現的差異，共有 16 題達顯著

表二 三組研究參與者各項基本能力

	ASD 組 n = 74	DD 組 n = 87	TD 組 n = 111	F / χ^2	p	效果值
生理年齡 (月)						
平均 (標準差)	24.18 (3.15)	22.25 (3.10)	24.02 (3.05)	10.39	< .001	.073
全距	18-29	18-29	18-29			
整體心理年齡當量 (月)						
平均 (標準差)	14.41 (3.83)	17.36 (3.88)	25.68 (4.68)	183.24	< .001	.569
全距	6.50-26.75	10.50-28.75	13.75-37.50			
父親教育年數						
平均 (標準差)	14.38 (2.71)	13.68 (2.66)	16.29 (1.87)	31.47	< .001	.191
全距	6-18	9-18	8-18			
母親教育年數						
平均 (標準差)	14.49 (2.33)	13.76 (2.87)	16.03 (1.37)	27.35	< .001	.169
全距	9-19	6-18	12-18			
性別 (人數) ^a						
男:女	65:9	62:25	53:58	33.36	< .001	.325
ADOS 分數 ^b						
平均 (標準差)	17.96 (2.68)	3.61 (1.70)		21.10	< .001	.914
全距	12-22	0-8				

^a 卡方檢定 (χ^2 test) · ^b 分數越高表示越符合自閉症特徵

差異（見表四）。ASD 組與 DD 組，共有 12 題達顯著差異，分別是：第 2 題、第 6 題、第 7 題、第 9 題、第 13 題到第 15 題、第 17 題、第 19 題，及第 21 題到第 23 題。ASD 組與 TD 組，共有 5 題未達顯著差異，分別是：第 1 題、第 3 題、第 4 題，第 12 題及第 18 題。三組幼兒在 M-CHAT 幾個指標分數的比較（見表五）。23 題總分，三組幼兒達顯著差異，ASD > DD > TD。關鍵 6 題、最佳 7 題及簡版 14 題的總分，三組幼兒達顯著差異，皆為 ASD > DD、TD。由於心理年齡會影響幼兒行為表現，且照顧者的教育程度可能影響對題目的解釋，進一步以共變數分析（analysis

of covariance, ANCOVA）控制心理年齡與母親的教育年數，再次檢驗 M-CHAT 幾個指標的差異，23 題總分、關鍵 6 題、最佳 7 題及簡版 14 題的總分，三組幼兒皆達顯著差異（ $p < .001$ ），最小差異法（least significant difference, LSD）進行事後檢定，結果皆為 ASD > DD、TD。

以訊號偵測理論決定 ASD 組與 DD 組及 ASD 組與 TD 組，M-CHAT 總分的最佳切截分數，結果發現 ASD 組與 DD 組，2.50-3.50 為相對較好的選擇；ASD 組與 TD 組，1.50-3.50 為相對較好的選擇，不同切截分數的篩檢正確率（見表六）。ASD 組與 DD 組以失敗 4

表三 修訂版學步期自閉症檢核表內部一致性

	每題與總分的相關	該題刪除時的 α
1、肢體接觸	.10	.82
2、社交動作	.46	.81
3、粗大動作一爬	.10	.82
4、捉迷藏	.32	.81
5、假裝遊戲	.47	.81
6、要求式手指指示	.57	.80
7、分享式手指指示	.58	.80
8、正確玩玩具	.28	.82
9、展示	.37	.81
10、眼神注視	.33	.81
11、聽覺敏感	-.06	.84
12、社會性微笑	.40	.81
13、模仿	.51	.80
14、叫名反應	.58	.80
15、跟隨手指指示	.66	.80
16、粗大動作一走	.00	.82
17、跟隨眼神注視	.58	.80
18、不尋常手勢動作	.03	.83
19、吸引注意	.55	.80
20、聽力問題	.43	.81
21、語言理解	.54	.80
22、眼神接觸不佳	.34	.81
23、社會參照	.32	.81

題作為切截分數，ASD 組與 TD 組以失敗 3 題作為切截分數，篩檢正確率最為理想。此外，ASD 組與 DD 組、ASD 組與 TD 組的偵測者運作特徵曲線下面積分別為 .79 與 .89。

本研究探討 M-CHAT 簡版的篩檢指標，包含 M-CHAT 作者提出的關鍵 6 題、最佳 7 題及 Wong (2018) 等人的簡版 14 題，探討 ASD 組與 DD 組、ASD 組與 TD 組的篩檢正確率 (見表七)，篩檢正確率最佳為簡版 14

題失敗 3 題。此外，偵測者運作特徵曲線下面積，ASD 組與 DD 組、ASD 組與 TD 組分別為 .81 與 .90。

以判別函數分析的結構矩陣，ASD 組與 DD 組及 ASD 組與 TD 組分別比較，找出區辨力相對較理想的題目，採余麗樺、郭乃文與潘秀琴 (1998) 的建議，相關係數 .32 (解釋量 10%) 作為標準，數值越高代表區辨力越好。ASD 組和 DD 組，有 9 題符合標準 (見

表四 三組研究參與者在修訂版學步期自閉症檢核表的表現差異

	ASD 組	DD 組	TD 組	ASD 組與 DD 組				ASD 組與 TD 組			
	<i>n</i> = 74	<i>n</i> = 87	<i>n</i> = 111	<i>p</i>	χ^2	<i>p</i>	效果值	χ^2	<i>p</i>	效果值	
1、肢體接觸	.15 (63:11)	.13 (76:11)	.10 (100:11)	.590	0.167	.428	.032	1.040	.214	.075	
2、社交動作	.34 (49:25)	.09 (79:8)	.10 (100:11)	< .001	14.836	< .001	.304	16.147	< .001	.295	
3、粗大動作 — 爬	.05 (70:4)	.01 (86:1)	.04 (107:4)	.314	1.200 ^a	.137	.122	0.049 ^a	.405	.043	
4、捉迷藏	.19 (60:14)	.18 (71:16)	.04 (107:4)	.014	0.932	.545	.007	5.822 ^a	.008	.195	
5、假裝遊戲	.34 (49:25)	.14 (75:12)	.04 (107:4)	< .001	9.028	.002	.237	28.354 ^a	< .001	.407	
6、要求式手指指示	.45 (41:33)	.05 (83:4)	.01 (110:1)	< .001	33.917 ^a	< .001	.474	53.633 ^a	< .001	.553	
7、分享式手指指示	.45 (41:33)	.08 (80:7)	.03 (108:3)	< .001	28.607	< .001	.422	47.079 ^a	< .001	.518	
8、正確玩玩具	.30 (52:22)	.20 (70:17)	.05 (106:5)	< .001	2.262	.094	.119	20.687 ^a	< .001	.350	
9、展示	.22 (58:16)	.02 (85:2)	.00 (111:0)	< .001	13.152 ^a	< .001	.306	23.607 ^a	< .001	.377	
10、眼神注視	.16 (62:12)	.09 (79:8)	.02 (109:2)	.002	1.812	.134	.106	11.208 ^a	< .001	.267	
11、聽覺敏感	.26 (55:19)	.28 (63:24)	.37 (70:41)	.193	0.075	.463	.022	2.569	< .001	.118	
12、社會性微笑	.14 (64:10)	.03 (84:3)	.00 (111:0)	< .001	4.816 ^a	.020	.184	13.324 ^a	.074	.293	
13、模仿	.43 (42:32)	.18 (71:16)	.02 (109:2)	< .001	11.803	.001	.271	48.107 ^a	< .001	.524	
14、叫名反應	.41 (44:30)	.01 (86:1)	.00 (111:0)	< .001	37.415 ^a	< .001	.498	50.767 ^a	< .001	.539	
15、跟隨手指指示	.43 (42:32)	.05 (83:4)	.01 (110:1)	< .001	32.210 ^a	< .001	.462	51.464 ^a	< .001	.542	
16、粗大動作 — 走	.00 (74:0)	.01 (86:1)	.00 (111:0)	.344	0.000 ^a	.540	.073	-	-	-	
17、跟隨眼神注視	.32 (50:24)	.09 (79:8)	.01 (110:1)	< .001	13.558	< .001	.290	35.121 ^a	< .001	.452	
18、不尋常手勢動作	.23 (57:17)	.23 (67:20)	.14 (95:16)	.215	0.000	.575	.000	2.219	.099	.110	
19、吸引注意	.49 (38:36)	.14 (75:12)	.05 (106:5)	< .001	23.217	< .001	.380	47.630 ^a	< .001	.521	
20、聽力問題	.39 (45:29)	.18 (71:16)	.10 (101:10)	< .001	8.589	.003	.231	24.308	< .001	.362	
21、語言理解	.34 (49:25)	.10 (78:9)	.02 (109:2)	< .001	13.188	< .001	.286	33.914 ^a	< .001	.444	
22、眼神接觸不佳	.46 (40:34)	.21 (69:18)	.10 (101:10)	< .001	11.665	.001	.269	33.418	< .001	.425	
23、社會參照	.27 (54:20)	.08 (80:7)	.06 (104:7)	< .001	10.322	.001	.253	15.294	< .001	.288	

失敗率 (是：否)

^a 萊茲連續校正 (Yates Continuity Correction)

表五 三組研究參與者修訂版學步期自閉症檢核表總分、關鍵 6 題、最佳 7 題及簡版 14 題總分

	ASD 組 <i>n</i> = 74	DD 組 <i>n</i> = 87	TD 組 <i>n</i> = 111	<i>F</i>	<i>p</i>	效果值
總分						
平均 (標準差)	6.80 (4.22)	2.59 (2.12)	1.29 (1.53)	96.70	< .001	.418
全距	0-17	0-10	0-10			
關鍵 6 題						
平均 (標準差)	2.27 (1.80)	0.44 (0.73)	0.15 (0.39)	98.16	< .001	.458
全距	0-6	0-3	0-2			
最佳 7 題						
平均 (標準差)	2.57 (2.01)	0.57 (0.80)	0.26 (.058)	89.74	< .001	.400
全距	0-7	0-3	0-3			
簡版 14 題						
平均 (標準差)	5.32 (3.71)	1.43 (1.45)	0.51 (0.96)	113.71	< .001	.403
全距	0-13	0-6	0-6			

表六 修訂版學步期自閉症檢核表總分不同切截分數篩檢正確率

切截分數	ASD 與 DD 組				ASD 與 TD 組			
	敏感度	特異度	PPV	NPV	敏感度	特異度	PPV	NPV
2	65/74 (.88)	30/87 (.34)	65/122 (.53)	30/39 (.77)	65/74 (.88)	78/111 (.70)	65/98 (.66)	78/87 (.90)
3	60/74 (.81)	49/87 (.56)	60/98 (.61)	49/63 (.78)	60/74 (.81)	92/111 (.83)	60/79 (.76)	92/106 (.87)
4	53/74 (.72)	63/87 (.72)	53/77 (.69)	63/84 (.75)	53/74 (.72)	102/111 (.92)	53/62 (.85)	102/123 (.83)

PPV= positive predictive value; NPV = negative predictive value

表七 修訂版學步期自閉症檢核表關鍵 6 題、最佳 7 題、簡版 14 題與診斷一致性訂版學步期自閉症檢核表總分不同切截分數篩檢正確率

切截分數	ASD 與 DD 組				ASD 與 TD 組			
	敏感度	特異度	PPV	NPV	敏感度	特異度	PPV	NPV
關鍵 6 題 ^a	46/74 (.62)	79/87 (.91)	46/54 (.85)	79/107 (.74)	46/74 (.62)	110/111 (.99)	46/47 (.98)	110/138 (.80)
最佳 7 題 ^b	47/74 (.64)	76/87 (.87)	47/58 (.81)	76/103 (.74)	47/74 (.64)	105/111 (.95)	47/53 (.89)	105/132 (.80)
簡版 14 題 ^c	53/74 (.72)	74/87 (.85)	53/66 (.80)	74/95 (.78)	53/74 (.72)	106/111 (.95)	53/58 (.91)	106/127 (.83)

PPPV= positive predictive value; NPV = negative predictive value

^a 切截分數 2 分, ^b 切截分數 2 分, ^c 切截分數 3 分

表八)；ASD 組和 TD 組，有 11 題符合標準 (見表九)。

依照判別函數分析的結果，以訊號偵測理論找出 ASD 組與 DD 組簡版 9 題的最佳切截分數，結果發現 0.50-2.50 為相對較好的選擇。以失敗 2 題作為切截分數，篩檢正確率最佳 (見表十)。此外，偵測者運作特徵曲線下面積為 .81。

依照判別函數分析的結果，以訊號偵測理論找出 ASD 組與 TD 組簡版 11 題的最佳切截分數，結果發現 0.50-2.50 為相對較好的選擇。以失敗 1 題作為切截分數，篩檢正確率最佳 (見表十一)。此外，偵測者運作特徵曲線下面積為 .91。

討論與結論

雖然 ASD 篩檢工具分成兩個層次，但同一份工具常會被用在不同場域進行篩檢。不過，篩檢指標應視對照組為發展問題或 TD 幼兒而有所差異。回顧學界的研究，同一份工具進行層次一或層次二篩檢時，很少有研究針對篩檢指標進行探究，這可能導致錯誤的篩檢結果。目前為止，臺灣 ASD 幼兒早期篩檢工具及效度探究都明顯不足，可能妨礙 ASD 幼兒的早期診斷與早期介入。因此，本研究選擇國際上常用且有中文版的 ASD 篩檢工具：M-CHAT，探究其在基

表八 修訂版學步期自閉症檢核表結構矩陣 (自閉症類群障礙症組與發展遲緩組)

題目	相關係數
14、叫名反應	.64
6、要求式手指指示	.60
15、跟隨手指指示	.58
7、分享式手指指示	.51
19、吸引注意	.45
9、展示	.36
2、社交動作	.35
17、跟隨眼神注視	.34
21、語言理解	.33
13、模仿	.31
22、眼神接觸不佳	.31
23、社會參照	.29
5、假裝遊戲	.27
20、聽力問題	.26
12、社會性微笑	.21
3、粗大動作一爬	.14
8、正確玩玩具	.13
10、眼神注視	.12
16、粗大動作一走	-.08
1、肢體接觸	.04
11、聽覺敏感	-.02
4、捉迷藏	.01
18、不尋常手勢動作	.00

表九 修訂版學步期自閉症檢核表結構矩陣 (自閉症類群障礙症組與一般發展組)

題目	相關係數
6、要求式手指指示	.53
15、跟隨手指指示	.52
14、叫名反應	.52
13、模仿	.50
19、吸引注意	.49
7、分享式手指指示	.49
17、跟隨眼神注視	.41
21、語言理解	.40
22、眼神接觸不佳	.38
5、假裝遊戲	.36
9、展示	.33
20、聽力問題	.31
8、正確玩玩具	.30
2、社交動作	.25
12、社會性微笑	.25
23、社會參照	.24
10、眼神注視	.22
4、捉迷藏	.16
11、聽覺敏感	-.10
18、不尋常手勢動作	.09
1、肢體接觸	.06
3、粗大動作一爬	.04

註：第 16 題因兩組幼兒填答結果皆為「是」，無相關係數。

層診所場域進行層次一篩檢，綜合醫院場域進行層次二篩檢，應該採用的篩檢指標及篩檢正確率。本研究的研究參與者共有 272 名 18-29 個月的幼兒，包括 74 名 ASD 幼兒、87 名 DD 幼兒及 111 名 TD 幼兒。

信度方面，M-CHAT 的內部一致性係數高於 .80，顯示其信度良好。然而，檢視 M-CHAT 每一項題目，有 5 題與總分的相關小於 .20，分別為第 1 題（肢體接觸）、第 3 題（粗大動作一爬）、第 11 題（聽覺敏感）、第 16 題（粗大動作一走）及第 18 題（不尋常手勢動作），這個結果與 Wong 等人（2018）的研究一致，支持 M-CHAT-R/F 刪減題目的作法。然而，在不同文化需要保留的題目可能不一樣，未來研究需要持續探究。檢視這 5 題，第 3 題與第 16 題為肢體動作及粗大動作，這些都不是 ASD 的核心症狀，相關低是合理的結果。第 1 題與第 11 題為感官敏感，第 18 題為怪異手勢動作，雖然是 ASD 的相關症狀，但可能因為年幼 ASD 較少出現（e.g., Rogers, 2009），導致家長可能較難察覺其異常，是造成相關較低

的一個原因。

效標關聯效度方面，臨床組（ASD 合併 DD 組）幼兒的 M-CHAT 總分與 ADOS 分數達中度顯著相關，支持 M-CHAT 是具效度的篩檢工具。三組幼兒在 M-CHAT 總分、關鍵 6 題、最佳 7 題，及簡版 14 題的總分，皆是 ASD 幼兒顯著高於 DD 與 TD 幼兒，亦支持 M-CHAT 是具備效度，可以用來進行 ASD 幼兒早期篩檢的工具。不過，有一個現象需要注意，M-CHAT 的 23 題中，ASD 幼兒沒有任何一題失敗率超過 50%，這個結果不同於 Robins 等人（2001）的研究，但與 Wong 等人（2018）的研究一致，顯示臺灣 ASD 幼兒的家長，很可能會低估小孩的行為症狀。臺灣的家長比較重視孩子的語言及認知發展，而較忽略社會互動的部分（Wang, Lee, Chen, & Hsu, 2012），這可能是家長低估小孩自閉症相關症狀的一個原因。此外，過去研究（Fountain, King, & Bearman, 2011; Hrdlicka et al., 2016）指出高社經地位的家庭，其 ASD 孩童被診斷出來的比例較高、年齡較小。本研究的 ASD 與 DD 幼兒皆來自

表十 修訂版學步期自閉症檢核表簡版 9 題與診斷一致性（自閉症類群障礙症組與發展遲緩組）

9 題版		ASD 與 DD 組		
切截分數	敏感度	特異度	PPV	NPV
1	57/74 (.77)	53/87 (.61)	57/91 (.63)	53/70 (.76)
2	53/74 (.72)	73/87 (.84)	53/67 (.79)	73/94 (.78)

表十一 修訂版學步期自閉症檢核表簡版 11 題與診斷一致性（自閉症類群障礙症組與一般發展組）

11 題版		ASD 與 TD 組		
切截分數	敏感度	特異度	PPV	NPV
1	66/74 (.89)	91/111 (.82)	66/86 (.77)	91/99 (.92)
2	54/74 (.73)	103/111 (.93)	54/62 (.87)	103/123 (.84)

於臺灣南部的鄉村地區，未來應該在臺北、臺中等都會區進行探究，探究 M-CHAT 在高社經地區的篩檢正確率，了解臺灣 ASD 幼兒盛行率偏低的可能原因。

由於 ASD 篩檢工具分成層次一與層次二。因此，篩檢正確率方面，將依據這個分類進行探討。篩檢正確率良好與否的指標，則是參照 Cicchetti、Volkmar、Klin 與 Showalter (1995) 的標準，小於 .70 為不佳 (poor)、.70-.79 為尚可 (fair)、.80-.89 為良好 (good)、大於 .90 為極佳 (excellent)。

(一) M-CHAT 篩檢 ASD 與 DD 幼兒的正確率

ASD 組與 DD 組在 M-CHAT 全部 23 題中，共 11 題未達顯著差異，分別是：第 1 題（肢體接觸）、第 3 題（粗大動作一爬）、第 4 題（捉迷藏）、第 5 題（假裝遊戲）、第 8 題（正確玩玩具）、第 10 題（眼神注視）、第 11 題（聽覺敏感）、第 12 題（社會性微笑）、第 16 題（粗大動作一走）、第 18 題（不尋常手勢動作）及第 20 題（聽力問題）。第 5 題與第 12 題，兩組幼兒沒有達到顯著差異的一個可能原因是本研究的標準較為嚴格 ($p < .002$)。另一個原因則是 DD 幼兒平均心理年齡約 17 個月，不少幼兒尚未發展出假裝遊戲的能力（吳進欽、朱慶琳、侯育銘、林姿伶、姜忠信，2014），未來研究需要擴大樣本持續探究。第 8 題（正確玩玩具）在 Robins 等人 (2014) 的 M-CHAT-R/F 中已刪除，顯示正確玩玩具的區辨力可能原本就不太理想。一個令人意外的結果是第 10 題（眼神注視），兩組幼兒沒有顯著差異，但是眼神接觸不佳卻是 ASD 幼兒早期重要的行為症狀。兩組幼兒眼神接觸沒有顯著差異的原因，可能是 M-CHAT 的題目是詢問

是否會注視 1-2 秒，家長會用有沒有眼神注視來回答。眼神注視常與互動溝通有關，但 M-CHAT 的題目沒有突顯這個特質。不過，M-CHAT-R/F 中，作者已將眼神注視的題目互動溝通特質呈現出來，未來的研究可以用 M-CHAT-R/F 探究臺灣家長對 ASD 幼兒眼神注視的判斷與解釋。

ASD 組與 DD 組篩檢正確率，以失敗 4 題作為切截分數，敏感度與特異度可以兼顧，正確率尚可，這個結果與 Wong 等人 (2018) 一致，但與 Robins 等人 (2001) 所建議的失敗 3 題不一致。檢視將 M-CHAT 用在臨床樣本的研究 (Eaves et al., 2006; Snow & Lecavalier, 2008; Salisbury et al., 2018)，若以 Robins 等人原本的建議，大多是特異度不佳，本研究亦是如此；失敗 3 題當作切截分數，容易出現過度診斷。然而，失敗 3 題可以讓 M-CHAT 敏感度提升為良好，優點是比較多的 ASD 幼兒被診斷出來，缺點是一些 DD 幼兒可能會被懷疑為 ASD。不過，DD 被懷疑成 ASD 的代價是幼兒可能會進行一些評估或介入。在臺灣尚缺乏特屬於 ASD 介入模式的情況下，若能留意與減少污名化的問題，以失敗 3 題作為切截分數亦是考慮的選項。

不同 M-CHAT 簡版篩檢 ASD 組與 DD 組的正確率，關鍵 6 題與最佳 7 題以失敗 2 題當作切截分數，皆為敏感度不佳 (.62-.64)，但特異度介於良好與極佳 (.87-.91)；簡版 14 題以失敗 3 題作為切截分數，則是敏感度尚可 (.72)，特異度良好 (.85)，結果與 Wong 等人 (2018) 的研究一致。若以本研究的簡版 9 題失敗 2 題作為切截分數，一樣是敏感度尚可 (.72)，特異度良好 (.84)，這個結果與 Wong 等人的簡版 14 題相近。綜合這些結果顯示 Robins 等人

(2001, 2010) 所建議的簡版，其對照組為 TD，用在臨床樣本需要謹慎。不同 M-CHAT 簡版的篩檢正確率，Wong 等人在 ASD 與 DD 幼兒中正確篩檢出 127 名，本研究正確篩檢出 126 名，Robins 等人 (2001) 的兩個指標正確篩檢出 123-125 名，皆高於全部 23 題失敗 4 題的 116 名，是在綜合醫院場域篩檢時可以考慮使用的方式。不過，考慮正確率指標的話，應該選擇 Wong 等人的簡版 14 題失敗 3 題。關鍵題目篩檢正確率略優於全部 23 題，這個結果不同於 Chlebowski 等人 (2013)，是文化差異或樣本屬性（一個綜合醫院場域，一個是基層診所場域）所導致，需要持續探究。

M-CHAT 用於綜合醫院場域的研究，23 題失敗 3 題，本研究的敏感度低於 Snow 與 Lecavalier (2008) 及 Wong 等人 (2018) 的研究；關鍵 6 題失敗 2 題，本研究的敏感度低於 Eaves 等人 (2006)、Snow 與 Lecavalier (2008) 及 Wong 等人 (2018) 的研究。不論是完整 23 題或不同簡版，本研究的篩檢正確率皆低於 Wong 等人 (2018) 的研究，特別是敏感度。檢視不同研究方法學上的差異，上述研究的參與者年紀比較大（48 個月以下兒童），本研究為 18-29 個月幼兒，ASD 幼兒年紀越小越不容易被診斷出來，這是本研究篩檢正確率較低的一個原因。過往的研究 (Kelly et al., 2019) 指出家長的教育程度越高，其 ASD 小孩被診斷出來的年紀越小。本研究 ASD 幼兒的家長平均教育年數約為 14 年，這樣的教育程度看似不低。然而，ASD 幼兒的家長教育程度低於 TD 幼兒的家長，且這一、二十年大學錄取率超過 95%；因此，ASD 幼兒的家長教育程度並不算太高，這可能也是本研究 ASD 幼兒篩檢正確率僅為尚可的一個原因。

相較於香港的研究 (Wong et al., 2004)，本研究結果的敏感度與特異度較低，兩篇研究存在一些方法學上的差異，這是 M-CHAT 篩檢正確率不同的原因。第一、Wong 等人 (2004) 的研究，乃以心理年齡約 24 個月為招募標準，ASD 幼兒生理年齡介於 26-86 個月。偏高的生理年齡篩檢正確率原本便會較為理想。再者，生理年齡高、心理年齡低，代表兒童障礙較為嚴重，家長也比較容易發現小孩的障礙。第二、Wong 等人 (2004) 的研究將 DD 與 TD 幼兒合併成一組對照組，這也會提高 M-CHAT 用在臨床樣本的敏感度與特異度。不過，一個令人意外的現象值得注意，本研究 ASD 幼兒的診斷依據為 DSM-5，其診斷標準較為嚴格，亦即自閉症特徵應該較為明顯，家長理應比較容易觀察到孩子社交互動的困難。然而，從篩檢正確率來看，家長似乎對於小孩自閉症症狀並未明顯的察覺，家長知覺與專家判斷間的差異，原因為何需要持續探究，這也彰顯需要提升低社經地區家長幼兒發展能力、人際互動能力相關知識的重要性。

目前臺灣以 DD 組作為對照組的 ASD 早期篩檢研究，吳進欽等人 (2014) 使用 CBS 篩檢 4 歲以下 ASD 兒童，敏感度與特異度分別是 .74 與 .73；吳進欽、侯育銘、朱慶琳與李羽涵 (2015) 以 CBS 篩檢 24-36 個月 ASD 幼兒，敏感度與特異度分別為 .78 與 .71，CBS 篩檢正確率屬於尚可。Wong 等人 (2018) 使用 M-CHAT 篩檢 18-47 個月 ASD 幼兒，敏感度為 .77，特異度為 .72，亦屬於尚可的程度，本研究結果亦顯示 M-CHAT 篩檢 ASD 幼兒的正確率屬於尚可。然而，T-STAT 篩檢 36 個月以下 ASD 幼兒，敏感度介於 .92- .95，特異度介於 .82- .95 (吳進欽等人, 2015；姜忠信等人, 2012)，

T-STAT 篩檢正確率優於家長問卷的 CBS 與 M-CHAT，顯示在臺灣於綜合醫院場域，使用家長問卷篩檢 ASD 幼兒的正確率有待提升。過去學界建議家長問卷後進行追蹤訪談澄清 (Kleinman et al., 2008)，家長問卷配合行為觀察工具 (吳進欽等人, 2015)，可以有效提升 ASD 篩檢正確率，未來的研究可以針對這兩種方式進行探究與比較，找出適合於臺灣於綜合醫院場域篩檢或診斷 ASD 幼兒的工作模式。

(二) M-CHAT 篩檢 ASD 與 TD 幼兒的正確率

ASD 組與 TD 組在 M-CHAT 的 23 題，共有 5 題未達顯著差異，分別是：第 1 題 (肢體接觸)、第 3 題 (粗大動作—爬)、第 4 題 (捉迷藏)、第 12 題 (社會性微笑) 及第 18 題 (不尋常手勢動作)，這些題目在 ASD 組與 DD 組也未達顯著差異，顯示這幾題不論在基層診所或綜合醫院場域進行篩檢，都不是具鑑別力的題目，使用 M-CHAT 時需要注意。這些题目的前三題，並不屬於 ASD 的核心症狀，缺乏鑑別力並不讓人意外。那麼社會性微笑與不尋常的手勢動作為何也缺乏鑑別力呢？社會性微笑是嬰幼兒 1 歲前常見的互動反應，在成人引發下嬰幼兒出現回應大人的交流與互動行為。共享式注意力 (joint attention) 的研究指出 ASD 幼兒回應大人所引發的互動行為之障礙較不嚴重 (e.g., Mundy & Newell, 2007)，原因是在大人引發幼兒注意有趣物體或事件時，ASD 幼兒注意到物體或事件，並給予回應，有趣的物體或事件會產生增強的效果。社會性微笑這種早期的社會互動能力，ASD 幼兒亦可能因為長期與重複的經驗而學習與建立起來，導致這個行為缺乏鑑別力。至於不尋常手勢動作，研究指出 ASD 幼兒雖然會出現同一

性 (sameness) 行為，但幼兒期同一性行為缺乏鑑別力 (e.g., Barton, Dumont-Mathieu, & Fein, 2012; Rogers, 2009)，本研究出現這個結果就不令人意外了。

ASD 組與 TD 組篩檢正確率，以失敗 3 題作為切截分數，敏感度與特異度皆在 .80 以上，屬於良好的程度，失敗 4 題作為切截分數，敏感度尚可 (.72)，特異度極佳 (.92)。若要兼顧敏感度與特異度，應以失敗 3 題作為切截分數。ASD 組與 TD 組的篩檢正確率明顯高於 ASD 組與 DD 組的篩檢正確率，顯示 M-CHAT 比較適合作為層次一的 ASD 篩檢工具，這與 Robins 等人 (2001) 發展 M-CHAT 的篩檢場合為基層診所場域一致。不過，需要注意的是失敗 4 題作為切截分數，在 ASD 與 TD 幼兒中可以正確篩檢出 155 名，但失敗 3 題作為切截分數只正確篩檢出 152 名。過去研究指出 ASD 篩檢工具在不同文化使用需要留意切截分數的調整 (Stewart & Lee, 2017)，雖然本研究發現失敗 3 題作為切截分數，可以同時兼顧敏感度與特異度，未來仍需要擴大樣本持續探究 M-CHAT 用在基層診所場域的篩檢指標。

不同 M-CHAT 簡版篩檢 ASD 與 TD 組的正確率，Robins 等人 (2001) 所建議的關鍵 6 題與最佳 7 題以失敗 2 題作為切截分數，敏感度與特異度分別為 .62-.64 與 .95-.99，特異度極佳但敏感度不佳。Wong 等人 (2018) 的簡版 14 題失敗 3 題作為切截分數，敏感度與特異度為 .72 與 .95，敏感度尚可，特異度極佳。若是以本研究的簡版 11 題失敗 1 題作為切截分數，敏感度與特異度皆屬於良好 (超過 .80)，失敗 2 題作為切截分數，敏感度尚可 (.73)，特異度極佳 (.93)。不同 M-CHAT 簡版的篩檢正確率，Wong 等人 (2018) 在 ASD 與 TD 幼兒中正確篩檢

出 159 名，本研究的簡版 11 題正確篩檢出 157 名，Robins 等人的兩個指標正確篩檢出 152-156 名。整體而言，雖然 M-CHAT 不同簡版的題目數存在差異，但篩檢正確率差異不大，且與全部 23 題的篩檢正確率差異也不大，在基層診所場域進行篩檢時，亦可以考慮使用簡版，Wong 等人（2018）的簡版 14 題失敗 3 題是相對較好的選擇。關鍵題目篩檢正確率略優於全部 23 題，這個結果和 Chlebowski 等人（2013）不一致，因為對照組是 TD，可以排除樣本屬性的影響，因此文化差異是較為可能的因素，如：照顧者對幼兒行為的解釋，未來需要持續探究。

臺灣以 TD 組作為對照組的 ASD 篩檢研究，Tsai 等人（2012）以 TASQ 篩檢 67 名 18-26 個月大的幼兒，其中 18 名為 ASD（15 名為中重度患者），敏感度與特異度分別為 1 與 .97，篩檢正確率極佳。阮麗敏與林玲伊（2015）以 ASRS-C 篩檢 165 名 24-71 個月的兒童，其中 30 名為 ASD，敏感度與特異度分別為 .97 與 .93，篩檢正確率極佳。本研究發現 M-CHAT 以失敗 3 題作為切截分數，敏感度與特異度分別是 .81 與 .83，篩檢正確率屬於良好。雖然 M-CHAT 篩檢正確率低於 TASQ 與 ASRS-C，不過 TASQ 的樣本數少且多數 ASD 幼兒為中重度患者，需要擴大樣本、納入症狀較輕微的 ASD 幼兒持續追蹤。ASRS-C 的 ASD 兒童一樣不多，且並非全部針對 30 個月以下幼兒，不符合早期篩檢的目的，篩檢 30 個月以下 ASD 幼兒的效度，需要再探究。M-CHAT 的篩檢正確率雖然低於 TASQ 與 ASRS-C，但還是相當理想，在基層診所場域進行 ASD 篩檢時，還是值得推廣的工具。

本研究使用 M-CHAT 篩檢 18-29 個月大的 ASD 幼兒，並分別以 DD 或 TD 幼兒作為

對照組，進行篩檢正確率探究。相較於國外研究（Yuen et al., 2018; Sánchez-García et al., 2019），若是綜合醫院場域，M-CHAT 在臺灣篩檢 ASD 幼兒的敏感度與特異度與國外相當；若是基層診所場域，M-CHAT 在臺灣篩檢 ASD 幼兒的敏感度略高、特異度略低。此外，本研究結果發現 M-CHAT 在臺灣篩檢 ASD 的正確率，若要兼顧敏感度與特異度，綜合醫院場域應該以 23 題失敗 4 題作為切截分數，基層診所場域應該以 23 題失敗 3 題作為切截分數。不過，在綜合醫院場域若是強調敏感度，則可以調整為 23 題失敗 3 題作為切截分數。此外，不同 M-CHAT 簡版，以 Wong 等人（2018）的 14 題失敗 3 題，篩檢正確率最為理想。整體而言，用 M-CHAT 全部 23 題與簡版進行篩檢，正確率結果差異不大，但簡版略優。本研究有以下限制，可以作為後續研究的參考：（1）本研究填寫者為主要照顧者，但絕大多數為母親，未來的研究可以比較父親與母親問卷填寫結果，比較兩者對幼兒行為觀察的一致性。（2）本研究的對象為臺灣南部地區，屬於鄉村地區。過去研究指出 ASD 診斷存在城鄉差異（Lung, Chiang, Lin, & Shu, 2017），未來的研究可以在中部及北部城市地區持續探究。（3）本研究為橫斷研究，以早期篩檢而言，樣本數還不夠大，特別是在基層診所場域。未來需要擴大樣本數，並以縱貫設計的方式，持續探討 M-CHAT 的篩檢正確率。

參考文獻

- 阮麗敏、林玲伊（2015）：「中文版泛自閉症篩檢量表」2~5 歲幼兒短版之信效度分析。測驗學刊，62（1），25-44。[Yuan, L. M., & Lin, L. Y. (2015). The reliability

- and validity of the Chinese version of autism spectrum rating scale: Short form (2-5 years). *Psychological Testing*, 62(1), 25-44.]
- 余麗樺、郭乃文、潘秀琴 (1998)：盧一尼神經心理測驗組中文修訂版的信效度研究。台灣精神醫學，12 (4)，365-378。[Yu, L. F., Guo, N. W., & Pan, H. C. (1998). Reliability and validity of the Chinese version of the Luria-Nebraska neuropsychological battery. *Taiwanese Journal of Psychiatry*, 12(4), 365-378.]
- 吳進欽、朱慶琳、侯育銘、林姿伶、姜忠信 (2014)：T-STAT 評估 18-24 個月自閉症類嬰幼兒效度探究。中華心理衛生學刊，27 (4)，621-644。[Wu, C. C., Chu, C. L., Hou, Y. M., Lin, T. L., & Chiang C. H. (2014). Validation of The screening tool for autism in two-year-olds Taiwan version (T-STAT) in children aged between 18 and 24 months. *Formosa Journal of Mental Health*, 27(4), 621-644.]
- 吳進欽、朱慶琳、侯育銘、姚淑芬 (2014)：克氏行為量表篩檢 4 歲以下自閉症類疾患兒童效度探究。中華心理衛生學刊，27 (1)，131-161。[Wu, C. C., Chu, C. L., Hou, Y. M., & Yao, S. F. (2014). Validity of the Clancy behavior scale for detecting autism spectrum disorders in children younger than 4 years old. *Formosa Journal of Mental Health*, 27(1), 131-161.]
- 吳進欽、姜忠信、侯育銘 (2013)：兩歲前自閉症類疾患嬰幼兒篩檢工具：文獻回顧。中華心理衛生學刊，26 (4)，549-568。[Wu, C. C., Chiang C. H., & Hou, Y. M. (2013). Screening tools for infants under two years old with autism spectrum disorders: A literature review. *Formosa Journal of Mental Health*, 26(4), 549-568.]
- 吳進欽、侯育銘、朱慶琳、李羽涵 (2015)：結合多重訊息來源篩檢 3 歲以下自閉症類群疾患幼兒。中華心理衛生學刊，28 (4)，605-630。[Wu, C. C., Hou, Y. M., Chu, C. L., & Lee, Y. H. (2015). Combining information from multiple sources to screen children with autism spectrum disorders under the age of 3 years. *Formosa Journal of Mental Health*, 28(4), 605-630.]
- 姜忠信、宋維村 (2005)：自閉症嬰幼兒的早期診斷：文獻回顧。臨床心理學刊，2 (1)，1-10。[Chiang C. H., & Soong W. T. (2005). Diagnosis of autism in infants and toddlers: A literature review. *Formosa Journal of Mental Health*, 2(1), 1-10.]
- 姜忠信、吳進欽、劉俊宏、侯育銘 (2012)：2-3 歲自閉症幼兒篩檢量表的發展：T-STAT 的初究。中華心理衛生學刊，25 (1)，135-154。[Chiang C. H., Wu, C. C., Liu J., H., & Hou, Y. M. (2012). Development of a screening tool for autism in two-year-olds, Taiwan version (T-STAT): A preliminary study. *Formosa Journal of Mental Health*, 25(1), 135-154.] doi:10.30074/FJMH.201203_25(1).0005
- 謝清芬、宋維村、徐澄清 (1983)：自閉症：克氏行為量表的效度與研究。中華民國神經精神醫學會會刊，9 (1)，17-27。[Hsieh, C. F., Soong, W. T., & Hsu, C. C. (1983). Autism: The validity of Clancy behavior scale. *Chinese Society of the Neurology and Psychiatry Bulletin*, 9(1),

- 17-27.]
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association. doi:10.1176/appi.books.9780890425596
- Barbaro, J. & Dissanayake, C. (2010). Prospective identification of autism spectrum disorders in infancy and toddlerhood using developmental surveillance: The social attention and communication study. *Journal of Developmental and Behavioral Pediatrics, 31*(5), 376-385. doi:10.1097/DBP.0b013e3181df7f3c
- Barbaro, J. & Dissanayake, C. (2013). Early markers of autism spectrum disorders in infants and toddlers prospectively identified in the social attention and communication study. *Autism, 17*(1), 64-86. doi:10.1177/1362361312442597
- Barbaro, J., & Dissanayake, C. (2017). Diagnostic stability of autism spectrum disorder in toddlers prospectively identified in a community-based setting: Behavioural characteristics and predictors of change over time. *Autism, 21*(7), 830-840. doi:10.1177/1362361316654084
- Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *The British Journal of Psychiatry, 161*(6), 839-843. doi:10.1192/bjp.161.6.839
- Barton, M. L., Dumont-Mathieu, T., & Fein, D. (2012). Screening young children for autism spectrum disorders in primary practice. *Journal of Autism and Developmental Disorders, 42*(6), 1165-1174. doi:10.1007/s10803-011-1343-5
- Bayley, N. (1969). *Bayley Scales of Infant Development*. San Antonio, TX: Psychological Corporation.
- Canal-Bedia, R., García-Primo, P., Martín-Cilleros, M. V., Santos-Borbujo, J., Guisuraga-Fernández, Z., Herráez-García, L.,...Posada-de La Paz, M. (2011). Modified checklist for autism in toddlers: Cross-cultural adaptation and validation in Spain. *Journal of Autism and Developmental Disorders, 41*(10), 1342-1351. doi:10.1007/s10803-010-1163-z
- Chakrabarti, S., & Fombonne, E. (2005). Pervasive developmental disorders in preschool children: Confirmation of high prevalence. *American Journal of Psychiatry, 162*(6), 1133-1141. doi:10.1176/appi.ajp.162.6.1133
- Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics, 131*(4), e1121-e1127. doi: 10.1542/peds.2012-1525
- Christensen, D. L., Bilder, D. A., Zahorodny, W., Pettygrove, S., Durkein, M. S., Fitzgerald, R. T., Eicem C., ... & Yeargin-Allsopp, M. (2016). Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network. *Journal of Developmental & Behavior Pediatrics, 37*(1), 1-8. doi:10.1097/

- DBP.0000000000000235
- Cicchetti, D. V., Volkmar, F., Klin, A., & Showalter, D. (1995). Diagnosing autism using ICD-10 criteria: A comparison of neural networks and standard multivariate procedures. *Child Neuropsychology*, *1*(1), 26-37. doi:10.1080/09297049508401340
- Dancey, C. P., & Reidy, J. (2007). *Statistics Without Maths for Psychology*. London: Prentice Hall.
- Eaves, L. C., Wingert, H., & Ho, H. H. (2006). Screening for autism: Agreement with diagnosis. *Autism*, *10* (3), 229-242. doi:10.1177/1362361306063288
- Fountain, C., King, M. D., & Bearman, P. S. (2011). Age of diagnosis for autism: Individual and community factors across 10 birth cohorts. *Journal of Epidemiology & Community Health*, *65*(6), 503-510. doi:10.1136/jech.2009.104588
- Guthrie, W., Swineford, L. B., Nottke, C., & Wetherby, A. M. (2013). Early diagnosis of autism spectrum disorder: Stability and change in clinical diagnosis and symptom presentation. *Journal of Child Psychology and Psychiatry*, *54*(5), 582-590. doi:10.1111/jcpp.12008
- Hrdlicka, M., Vacova, M., Oslejskova, H., Gondzova, V., Vadlejchova, I., Kocourkova, J., Koutek, J., & Dudova, I. (2016). Age at diagnosis of autism spectrum disorders: Is there an association with socioeconomic status and family self-education about autism? *Neuropsychiatric Disease and Treatment*, *12*, 1639-1644. doi:10.2147/NDT.S107239
- Inada, N., Koyama, T., Inokuchi, E., Kuroda, M., & Kamio, Y. (2011). Reliability and validity of the Japanese version of the modified checklist for autism in toddlers (M-CHAT). *Research in Autism Spectrum Disorders*, *5*(1), 330-336. doi:10.1016/j.rasd.2010.04.016
- Kamio, Y., Inada, N., Koyama, T., Inokuchi, E., Tsuchiya, K., & Kuroda, M. (2014). Effectiveness of using the modified checklist for autism in toddlers in two-stage screening of autism spectrum disorder at the 18-month health check-up in Japan. *Journal of Autism and Developmental Disorders*, *44*(1), 194-203. doi:10.1007/s10803-013-1864-1
- Kelly, B., Williams, S., Collins, S., Mushtaq, F., Mon-Williams, M., Wright, B., ... & Wright, J. (2019). The association between socioeconomic status and autism diagnosis in the United Kingdom for children aged 5–8 years of age: Findings from the Born in Bradford cohort. *Autism*, *23*(1), 131-140. doi:10.1177/1362361317733182
- Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H. C., Esser, E. L., Wilson, L. B., Rosenthal, M. A., Sutera, S., Verbalis A. D., Barton, M., Hodgson, S., Green, J., Dumont-Mathieu, T., Volkmar, F., Chawarska, K., Klin, A., Fein, D. (2008). The modified checklist for autism in toddlers: A follow-up study investigating the early detection of autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(5), 827-839. doi:10.1007/s10803-007-0450-9
- Koh, H. C., Lim, S. H., Chan, G. J., Lin, M. B., Lim, H. H., Choo, S. H. T., & Magiati, I. (2014). The clinical utility of the modified checklist for autism in toddlers with high

- risk 18-48 month old children in Singapore. *Journal of Autism and Developmental Disorders*, 44(2), 405-416. doi:10.1007/s10803-013-1880-1
- Kondolot, M., Özmert, E. N., Öztop, D. B., Mazıcıolu, M. M., Gümü, H., & Elmalı, F. (2016). The modified checklist for autism in Turkish toddlers: A different cultural adaptation sample. *Research in Autism Spectrum Disorders*, 21, 121-127. doi:10.1016/j.rasd.2015.10.006
- Lai, D. C., Tseng, Y. C., Hou, Y. M., & Guo, H. R. (2012). Gender and geographic differences in the prevalence of autism spectrum disorders in children: Analysis of data from the national disability registry of Taiwan. *Research in Developmental Disabilities*, 33(3), 909-915. doi:10.1016/j.ridd.2011.12.015
- Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule Second Edition (ADOS-2) Manual (Part 1): Modules 1-4*. Torrance, CA: Western Psychological Services.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24(5), 659-685. doi:10.1007/BF02172145
- Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule (ADOS)*. LA: Western Psychological Services.
- Lung, F. W., Chiang, T. L., Lin, S. J., & Shu, B. C. (2011). Autism-risk screening in the first 3 years of life in Taiwan Birth Cohort Pilot Study. *Research in Autism Spectrum Disorders*, 5(4), 1385-1389. doi:10.1016/j.rasd.2011.01.024
- Lung, F. W., Chiang, T. L., Lin, S. J., & Shu, B. C. (2017). Urban and education disparity for autism spectrum disorders in Taiwan Birth Cohort Study. *Journal of Autism and Developmental Disorders*, 47(3), 599-606. doi:10.1007/s10803-016-2980-5
- Mohamed, F., Zaky, E., Youssef, A., Elhossiny, R., Zahra, S., Khalaf, R., . . . Abd-Elhakim, R. (2016). Screening of Egyptian toddlers for autism spectrum disorder using an Arabic validated version of M-CHAT; report of a community-based study (Stage I). *European Psychiatry*, 34, 43-48. doi:10.1016/j.eurpsy.2016.01.2421
- Mullen, E. M. (1995). *Mullen Scales of Early Learning: AGS*. Circle Pines, MN: American Guidance Service.
- Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, 16(5), 269-274. doi:10.1111/j.1467-8721.2007.00518.x
- Osterling, J., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology*, 14(2), 239-251. doi:10.1017/S0954579402002031
- Perera, H., Wijewardena, K., & Aluthwelage, R. (2009). Screening of 18-24-month-

- old children for autism in a semi-urban community in Sri Lanka. *Journal of Tropical Pediatrics*, 55(6), 402-405. doi:10.1093/tropej/fmp031
- Robins, D., Pandey, J., Chlebowski, C., Carr, K., Zaj, J., Arroyo, M., . . . Fein, D. (2010). M-CHAT Best7: A new scoring algorithm improves positive predictive power of the M-CHAT. *Paper Presented at the International Meeting for Autism Research (IMFAR)*.
- Robins, D. L. (2008). Screening for autism spectrum disorders in primary care settings. *Autism*, 12(5), 537-556. doi:10.1177/1362361308094502
- Robins, D. L., Casagrande, K., Barton, M., Chen, C. M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics*, 133(1), 37-45. doi:10.1542/peds.2013-1813
- Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 31(2), 131-144. doi:10.1023/A:1010738829569
- Rogers, S. J. (2009). What are infant siblings teaching us about autism in infancy? *Autism Research*, 2(3), 125-137. doi:10.1002/aur.81
- Rogers, S. J., Hepburn, S. L., Stackhouse, T., & Wehner, E. (2003). Imitation performance in toddlers with autism and those with other developmental disorders. *Journal of Child Psychology and Psychiatry*, 44(5), 763-781. doi:10.1111/1469-7610.00162
- Salisbury, L. A., Nyce, J. D., Hannum, C. D., Sheldrick, R. C., & Perrin, E. C. (2018). Sensitivity and specificity of 2 autism screeners among referred children between 16 and 48 months of age. *Journal of Developmental & Behavioral Pediatrics*, 39(3), 254-258. doi:10.1097/DBP.0000000000000537
- Samadi, S. A., & McConkey, R. (2015). Screening for autism in Iranian preschoolers: Contrasting M-CHAT and a scale developed in Iran. *Journal of Autism and Developmental Disorders*, 45(9), 2908-2916. doi:10.1007/s10803-015-2454-1
- Sánchez-García, A. B., Galindo-Villardón, P., Nieto-Librero, A. B., Martín-Rodero, H., & Robins, D. L. (2019). Toddler screening for autism spectrum disorder: A meta-analysis of diagnostic accuracy. *Journal of Autism and Developmental Disorders*, 49(5), 1837-1852. doi:10.1007/s10803-018-03865-2
- Seif Eldin, A., Habib, D., Noufal, A., Farrag, S., Bazaid, K., Al-Sharbati, M., . . . Gaddour, N. (2008). Use of M-CHAT for a multinational screening of young children with autism in the Arab countries. *International Review of Psychiatry*, 20(3), 281-289. doi:10.1080/09540260801990324
- Seung, H., Ji, J., Kim, S. J., Sung, I., Youn, Y. A., Hong, G., . . . Youm, H. K. (2015). Examination of the Korean modified

- checklist of autism in toddlers: Item response theory. *Journal of Autism and Developmental Disorders*, 45(9), 2744-2757. doi:10.1007/s10803-015-2439-0
- Snow, A. V., & Lecavalier, L. J. A. (2008). Sensitivity and specificity of the modified checklist for autism in toddlers and the social communication questionnaire in preschoolers suspected of having pervasive developmental disorders. *Autism*, 12(6), 627-644. doi:10.1177/1362361308097116
- Srisinghasongkram, P., Pruksananonda, C., & Chonchaiya, W. (2016). Two-step screening of the modified checklist for autism in toddlers in Thai children with language delay and typically developing children. *Journal of Autism and Developmental Disorders*, 46(10), 3317-3329. doi:10.1007/s10803-016-2876-4
- Stewart, L. A., & Lee, L. C. (2017). Screening for autism spectrum disorder in low-and middle-income countries: A systematic review. *Autism*, 21(5), 527-539. doi:10.1177/1362361316677025
- Sturner, R., Howard, B., Bergmann, P., Stewart, L., & Afarian, T. E. (2017). Comparison of autism screening in younger and older toddlers. *Journal of Autism and Developmental Disorders*, 47(10), 3180-3188. doi:10.1007/s10803-017-3230-1
- Toh, T. H., Tan, V. W. Y., Lau, P. S. T., & Kiyu, A. (2018). Accuracy of modified checklist for autism in toddlers (M-CHAT) in detecting autism and other developmental disorders in community clinics. *Journal of Autism and Developmental Disorders*, 48(1), 28-35. doi:10.1007/s10803-017-3287-x
- Topçu, S., Ulukol, B., Öner, Ö., İmek Orhon, F., & Bakan, S. (2018). Comparison of tidos with M-CHAT for screening autism spectrum disorder. *Psychiatry and Clinical Psychopharmacology*, 28(4), 416-422. doi:10.1080/24750573.2017.1422682
- Tsai, W. C., Soong, W. T., & Shyu, Y. I. (2012). Toddler autism screening questionnaire: Development and potential clinical validity. *Autism*, 16(4), 340-349. doi:10.1177/13623613111429694
- Veness, C., Prior, M., Bavin, E., Eadie, P., Cini, E., & Reilly, S. (2012). Early indicators of autism spectrum disorders at 12 and 24 months of age: A prospective, longitudinal comparative study. *Autism*, 16(2), 163-177. doi:10.1177/13623613111399936
- Wallis, K. E., & PintoMartin, J. (2008). The challenge of screening for autism spectrum disorder in a culturally diverse society. *Acta Paediatrica*, 97(5), 539-540. doi:10.1111/j.1651-2227.2008.00720.x
- Wang, J., Lee, L. C., Chen, Y.S., & Hsu, J. W. (2012). Assessing autistic traits in a Taiwan preschool population: Cross-cultural validation of the social responsiveness scale (SRS). *Journal of Autism and Developmental Disorders*, 42(11), 2450-2459. doi:10.1007/s10803-012-1499-7
- Wong, V., Hui, L. H. S., Lee, W. C., Leung, L. S. J., Ho, P. K. P., Lau, W. L. C., . . . Chung, B. (2004). A modified screening tool for autism (checklist for autism in toddlers [CHAT-23]) for Chinese children. *Pediatrics*, 114(2), e166-e176. doi:10.1542/

- peds.114.2.e166
- Wong, Y. S., Yang, C. C., Stewart, L., Chiang, C. H., Wu, C. C., & Iao, L. S. (2018). Use of the Chinese version modified checklist for autism in toddlers in a high-risk sample in Taiwan. *Research in Autism Spectrum Disorders, 49*, 56-64. doi:10.1016/j.rasd.2018.01.010
- Yuen, T., Penner, M., Carter, M. T., Szatmari, P., & Ungar, W. J. (2018). Assessing the accuracy of the modified checklist for autism in toddlers: A systematic review and metaanalysis. *Developmental Medicine & Child Neurology, 60*(11), 1093-1100. doi:10.1111/dmcn.13964

收稿日期：2019.06.12

接受日期：2019.10.02

Detection of Autism Spectrum Disorder in Different Settings: Accuracy of the Modified Checklist for Autism in Toddlers

Ying-Chia Hung

Postgraduate Student,
Dept. of Psychology,
Kaohsiung Medical University

I-Chen Chen

Clinical Psychologist,
Changhua Christian Children's
Hospital Center for Child
Development

Chin-Chin Wu

Professor/Adjunct Research Fellow,
Dept. of Psychology/Dept. of
Medical Research,
Hospital Kaohsiung Medical
University

ABSTRACT

Purpose: The Modified Checklist for Autism in Toddlers (M-CHAT) is a well-researched screening tool that was originally developed for detecting autism spectrum disorder (ASD) in toddlers. However, few studies have examined the accuracy of the M-CHAT for detecting ASD in the Taiwanese population. Thus, this study examined the accuracy of the M-CHAT for detecting ASD in Taiwanese toddlers in different settings (i.e., primary clinics and general hospitals). **Methods:** The M-CHAT comprises 23 questions with yes/no responses, and it is designed to detect ASD in toddlers aged 16-30 months. In this study, the M-CHAT was administered to 272 caregivers of toddlers aged 18-29 months. The 272 caregivers comprised those of 74 toddlers with ASD, 87 toddlers with developmental delay (DD), and 111 toddlers who were typically developing (TD). The chi-square test was used to examine the failure rate of each M-CHAT item among the ASD, DD, and TD groups. Analysis of variance was also used to examine the sensitivity and specificity of the total scores of the 23 M-CHAT items and the various short forms among the three groups of participants. In addition, signal detection theory was used to determine the optimal cutoffs of the M-CHAT, and discriminant function analysis was conducted to obtain the structure matrix for deciding the critical items. **Results/Findings:** When the M-CHAT was used for discriminating between the

ASD and DD groups in general hospitals, sensitivity of .72 and specificity of .72 were obtained. The cutoff was the failure of any 4 of the 23 M-CHAT items. For the different short forms of the M-CHAT, sensitivities of .62- .72 and specificities of .85- .91 were obtained; “Brief 14” was the best short form of the M-CHAT. When the M-CHAT was used for discriminating between ASD and TD groups in primary clinics, sensitivity of .81 and specificity of .83 were obtained. The cutoff was the failure of any 3 of the 23 M-CHAT items. For different short forms of the M-CHAT, sensitivities of .62- .72 and specificities of .95- .99 were obtained; “Brief 14” was the best short form of the M-CHAT.

Conclusions/Implications: The accuracy of the M-CHAT was adequate in general hospital settings and favorable in primary clinic settings. The study results suggest that the M-CHAT can be used as a screening tool for detecting ASD in toddlers in Taiwan, especially in primary clinics. However, different cutoffs in the form of different M-CHAT scores should be used depending on setting (primary clinic or general hospital).

Keywords: autism spectrum disorder, Modified Checklist for Autism in Toddlers (M-CHAT), specificity, sensitivity

